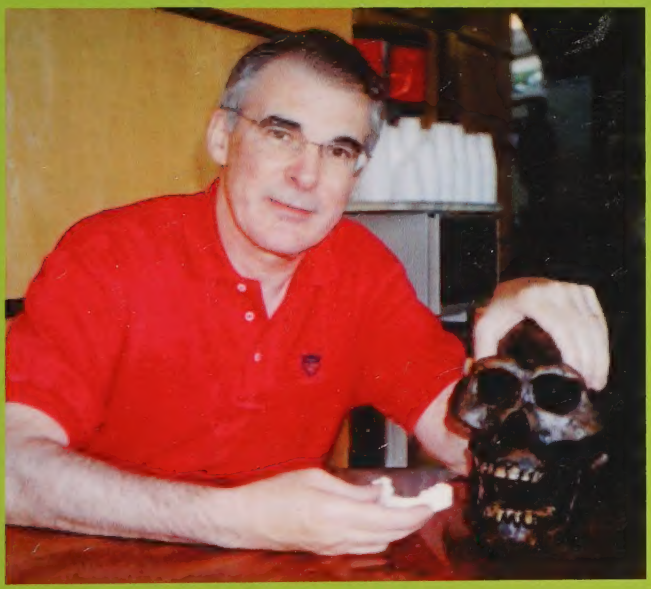
THE LIFE HISTORY APPROACH TO HUMAN DIFFERENCES

A Tribute to J. Philippe Rushton



Helmuth Nyborg
Editor

Digitized by the Internet Archive
in 2023 with funding from
Kahle/Austin Foundation

https://archive.org/details/isbn_9780993000119

THE LIFE HISTORY APPROACH TO HUMAN DIFFERENCES

A Tribute to J. Philippe Rushton

Helmuth Nyborg (Editor)

UI

Ulster Institute for Social Research

Published 2015 by
Ulster Institute for Social Research
London NW10 5TH, UK

ISBN 9780993000119

The articles in this book were first published in
Personality and Individual Differences,
Volume 55, Issue 3, (July 2013), pp. 199-342.
All necessary permissions have been obtained.

Copyright ©2015 Helmuth Nyborg

All rights reserved

Printed in Great Britain

Contents

Part 1: The Man

Obituary 7
Helmuth Nyborg

In conversation with J. Phillippe Rushton 17
Helmuth Nyborg

Part 2: The Book

Introduction 39

Part 3: Intelligence

Chapter 1. 47
Arthur R. Jensen: Rushton's Contributions to the Study of Mental Ability.

Chapter 2. 65
Linda Gottfredson: Resolute Ignorance on Race and Rushton.

Chapter 3. 85
Jan te Nijenhuis: The Flynn Effect, Group Differences and *g* loadings.

Chapter 4. 103
Heiner Rindermann: African Cognitive Ability: Research, Results, Divergences and Recommendations.

Part 4: Personality and development

Chapter 5. 179
Paul Irwing: The General Factor of Personality: Substance or Artefact?

Chapter 6. 213
Donald Templer: Rushton: The Great Theoretician and his Contribution to Personality.

Chapter 7. 233
Yoon-Mi Hur: J.P. Rushton's Contribution to the Study of Altruism.

Chapter 8. 247
Aurelio J. Figueredo, Tomás Cabeza de Baca & Michael A. Woodley: The Measurement of Human Life History Strategy.

Chapter 9. 267
Frank Salter & Henry Harpending: J.P. Rushton's Theory of Ethnic Nepotism.

Chapter 10. 287
Richard Lynn: An Examination of Rushton's Theory of Differences in Penis Length and Circumference and *r-K* History Theory in 113 Populations.

Chapter 11. 307
Helmuth Nyborg: Migratory Selection for Inversely Related Covariant T- and IQ-Nexus Traits: Testing the IQ/T-Geo-Climatic-Origin Theory by the General Trait Covariance Model.

Chapter 12. 327
Gerhard Meisenberg & Michael Woodley: Global Behavioural Variation: A Test of Differential-K.

Chapter 13. 347
Michael Woodley & Gerhard Meisenberg: A Jensen Effect on Dysgenic Fertility: An Analysis involving the National Longitudinal Study of Youth.

Index 361

Part 1: The Man

Obituary

J. Philippe Rushton: Eminent scientist, hero, and friend died 2nd October 2012.

Psychology professor J. Philippe Rushton died October 2nd 2012 from Addison's Disease, a condition compromising the immune system. This was also the time when Psychology, Behavior Genetics, and Evolutionary Theory lost one of their prominent figures, and we lost a supportive friend in the fight against political correctness.

Phil Rushton was born in England in 1943, emigrated to Canada in 1974, and became full professor of psychology in 1985 at the University of Western Ontario, Canada. Here he combined - over most of his professional career - the unusual skills of an uncompromising scientist with the behavior of a true English gentleman.

Rushton's personal background and professional career were recently laid out in a detailed interview with him (Nyborg, 2012, and Part I in this book), so I will not duplicate it in this obituary. I rather want to address a problem intimately associated with Rushton's particular research interests - evolutionary theory, genetic similarity theory, *r/K*-theory, life history theory, sex differences, brain size differences, and altruism (e.g. Rushton, 1995).

The problem anybody with such research interests constantly run into is that many social-constructivist academics not only find the biological parts of these areas particularly troubling, yes, even offensive, when applied to humans, but also that they think that their moralizing gives them carte blanche to viciously attack all such interested, and to isolate them within, or exclude

them from, academia. This sounds like an Orwellian world rather than a de facto characterization of the current conditions for research in biologically inspired behavioral science. Even if Rushton spent a lifetime collecting a mountain of relevant data, despite testing falsifiable hypotheses, and despite discussing alternative theories, he came out as a bad scientist everybody – scientist or layman - could dip in acid.

The historical background for this aversion in Behaviorist circles and in Social Constructivist thinking, which took over much of behavioral science ever since the 1930es, is that the critiques see Man as essentially a blank slate, upon which experiences writes. Their idea has for too long dominated philosophical and public thinking (Nyborg, 2003; Pinker, 2002), and has become almost a mantra in essential quarters of the behavioral sciences, to be repeated 5 times every day before breakfast. When Rushton again and again challenged the mantra he inveterated Voltaire's famous diction: "*Il est dangereux d'avoir raison dans des choses où des hommes accrédités ont tort*" (It is dangerous to be right in matters where established Men (or authorities) are wrong).

Rushton challenged the mantra when he sympathized with Charles Darwin's evolutionary theory, applied to animals including Man - so much, that he established the Charles Darwin Research Institute. He sympathized with socio-biologist E.O. Wilson and with Bill Hamilton – so much that he extended their *r-K* and altruism theories. He sympathized with (in)famous luminaries like Francis Galton, Charles Spearman, Raymond Cattell, Hans Eysenck, Arthur Jensen, and Richard Lynn – so much that he not only joined, but eventually extended the scope of the (in)famous *London School of Differential Psychology*. Still worse, Phil appreciated the brain power, incisiveness, and support from outside controversial scientists and co-heroes like Nobelists James Watson and William Shockley. Like them, Phil found out the hard way how dangerous it is to line up with currently unpopular people, ideas, or theories, which established Men are either ignorant of, or find them despicable.

Of course, Phil was not a first. Newton was no stranger to controversy, but his times nevertheless encouraged him to exclaim: *If I have seen further it is by standing on ye sholders of Giants* [sic]. Much later physics Nobelist Murray Gel-Man - who among many other things "discovered" the Quark – was allowed to jokingly paraphrase Newton's statement: *If I have seen further than others, it is because I am surrounded by dwarfs* (Gelfer, 2009). In lieu of the ill-intentioned full-scale controversy Phil constantly ran into, he might have asserted: *If I have seen further, it is because I follow forbidden data, and fed the hordes of academics throwing eggs at me, with still worse data.* Like Giordano Bruno, Galileo, and many others, Phil found out the hard way what Voltaire meant: They all spoke from personal experience. Like the others, Phil became a victim of the prejudices and deep ignorance of his time. There obviously is a price to be paid for this, even if it differs over time. After Freud escaped Germany and the Nazis burned his books ("*Ich ergebe zum Feur …*"), he told the press that this certainly was progress: Previously, they would have burned me, too! Canadian authorities did not burn Phil, but they confiscated one of his books; the Canadian police investigated him for accusation of "Hate speech", and the then Premier of Ontario phoned the university to call for his resignation. Do we spot similarities here? Shame on those established Canadian authorities!

Phil always responded to threats in a remarkably well-behaved way, as the true academic he was. A video-intermezzo between him and geneticist David Suzuki back in 1989 illustrates this prominent personality trait well. Here, Phil first composedly presents the relevant scientific data, followed by geneticist David Suzuki and others, who bring out a bravado of moral outrage over his head, in addition to insults and ad hominem attacks galore (Unbelievable, but see for yourself, and also read an informed Obituary at: http://www.amren.com/news/2012/10/phil-rushton-has-died/). When Suzuki aggressively called for Phil to be fired, Phil calmly responded: "That is not a scientific argument." When accused of being a racist, Phil answered: I am an academic". When an Afro-American attendant asked if Phil saw him as inferior, Phil

answered: We are different. Phil always stressed that moral incentives doesn't add to science. As could be expected, Suzuki also brought up the legendary Arthur Jensen from Berkeley University - whom Phil admired, was inspired of, and worked together with - accusing both of doing racist research. While watching the video, I recalled, that Art once sent me a couple of videotapes documenting similar vitriolic academic and public attacks on him, after he published his invited authoritative 1969 paper: "*How much can we boost IQ and scholastic achievement*". Not much, he concluded, based on more than 100 pages of studies and data. This was also too much for most established Men, so they deemed the data and theory flawed, the man wicked, and all failing their moral high ground standards. Well-brought-up, decent, established people also wanted to fire him.

Not being able to fire them, Phil's university restricted for some time his lecturing to showing pre-taped videos in private rooms. Art had to flee his private home for some time and was not allowed to enter campus without body-guards. This is not the Spanish Inquisition but life at modern universities. O Tempora. O Mores.

Phil and Art retorted by joining forces and summarizing ten lines of evidence from thirty years of research on race differences in cognitive ability (Rushton & Jensen, 2005). As expected, the impressive compilation of data, which they used to decide among contrasting hypotheses, did not impress their morally outraged colleagues.

Having had the privilege of enjoying Phil's friendship, I inadvertently became involved in a couple of his many unpleasant experiences. Once, as Phil and I paused at a conference, a well-dressed man approached us. He instantly began to see the accusations of racism, then folded his fists and punched Phil hard on the chest. As I went to his defense, Phil gently took my arm and said in a soft voice: No, No, Helmuth, the man is defenseless – he has no data. The attacker walked dumbfounded away. The same night our behavior genetics conference offered a privately

sponsored symposium where Phil, Art Jensen, the eminent behavior geneticist John C. Loehlin, among others, were viciously attacked by a great number of livid colleagues. I attended, as did John, but Art and Phil decided to have a cup of coffee elsewhere. Perhaps they remained assured that, despite all the noises, the earth is still not at the center and moves; species evolve while doing so, and altruism re-appears despite the constant loss of altruistic heroes. However much noise the critiques make, intelligence can still be estimated with fair approximation, is largely inherited (60-80%), and disposes for stable and consequential individual, group, and educational, economic, and geographical differences. Neither moral outcry nor low punches change any of that.

Unfortunately, the rule for political correct critique seems to be that the more convincing the evidence Phil and others amass, the worse will the messenger be treated (Nyborg, 2011). This may, in fact explain why Phil's academic style and polite personality didn't do him much good over the years. Acknowledging that Phil was exceptionally good at presenting hard data, critiques seem to have realized that the most effective way of downplaying his message was to systematically have a go for the man, while at the same time ignoring the mountain of highly relevant Behavior Genetic evidence emanating from countless family, twin-, and adoption-studies, genetic-distance studies, and from molecular biology and brain-scanning studies. New and better data will just electrify opponents even more in their anti-rational high-horse world.

Phil's career therefore raises a difficult question: If solid data and a classical academic approach do not suffice to avert the threat to Phil's research in particular and to differential psychology and evolutionary theory in general, then what will?

Not being as quite gentle as Phil, I suggest we now have to more visibly and systematically expose the critiques as the collaborators in the greatest collective scientific fraud of the 20th century, they are (Gottfredson, 1994, 2000; Nyborg, 2003; 2011). Each time they attack biologically inspired research, we have to ask them why they always leave out all the relevant behavior genetics

evidence, which a priori would have brought their empirical and/or moral position in question. Another line of action is to lay bare the elusive concepts, ideology, phraseology, and consequences of morally guided or left-oriented political critique (Gottfredson, 1994, 2000; Gross & Levitt, 1994/1998; Segerstråle, 2000), and to ask them what good comes out of a blatant confusion about ours "what is", and theirs "what ought to be". A third is to expose those who accuse others of fiddling with data on race differences in brain size, while all the time themselves fiddling. A good example of this is the study by Lewis et al. (2011). They re-analyzed the large skull sample, which paleontologist Steven Jay Gould from Harvard University originally accused the late Samuel Morton of fiddling with (Gould, 1981/1996) - with negative consequences for Phil's research on race differences in brain size (Rushton, 2010). The re-analysis proved Morton right and Gould wrong. Another relevant study is Rushton and Ankney's (2000) own re-examination of questionable brain size data. The exposure of claimants of scientific fraud, made in the name of anti-racism or political correctness, has made life much easier for a lot of us working under the constant scrutiny of sympathizers of Gould's (1981/1996) highly influential book *The Mismeasure of Man* (Nyborg, 2003, Rushton, 1997).

To sum up, I find that the life and science of Phil Rushton tell us two important lessons: Evolutionary theory, differential psychology, and behavior genetics are not only sciences but, unfortunately, also academic and public war zones. Progress in these areas accordingly requires a more active and determined counteraction from those of us, who aren't yet scared away from collecting even better biologically based behavioral data, and who are interested in analyzing how they relate to important environmental factors in guiding development and behavior, and producing similarities and differences.

Phil was an inspiration for us all to finally stand up and carry out our duty as scientists: To re-awaken the many fearfully silenced academic chickens, who well dare whispers in our ears that they agree, but who, in the presence of colleagues, at conferences, or

near a journalist, don't dare open their fine little mouths and tell the unpleasant truths they so fearfully carry within themselves. They have to be told to free themselves from the shameful shadowy halls of politically correct academia, and asked to start acting like real scientists with a spine.

Obviously, critique is essential for science, but it has to be informed and fair. The lack of both these latter aspects made Phil's life and professional career flip between greatness and tragedy. Our duty is accordingly to better inform established Men - to make it clear to them that there is a desperate need for radical changes in the bad climate for biologically attentive behavioral sciences towards less moralizing and more science. Phil's case makes it clear that not only people in general, but politicians and the leading elites in particular, need to be told in clear and no uncertain terms that there are valid scientific alternatives to supercilious blank slate ideas and condescending social constructivism. This is where science and society should now meet, and science is indeed needed for established people to enable them to make informed decisions about the future of individuals, groups, and whole nations.

Phil, the gentlemen and lone scientist - ironically so interested in altruism - tried hard. For this we ought to be eternally grateful. We now have to try even harder. It this endeavor we certainly will miss him as our beacon for changes for the betterment of behavioral science, policy, and an altruistic humanity.

References

Genfer, A. (10 September, 2009). Wilson vs Watson: The blessing of great enemies. *New Scientist.*

Gottfredson, L. (1994). Egalitarian fiction and collective fraud. *Society, 31*(3), Whole No. 209.

Gottfredson, L. (2000). Equal potential: a collective fraud". *Society, 37*(5), vii/viii.

Gould, S.J. (1981/1996). The Mismeasure of Man. New York: W.W. Norton/Hammondsworth, U.K.: Penguin Books.

Gross, P.R., & Levitt, N. (1994/1998). Higher superstition: The academic left and its quarrels with science. Baltimore: Johns Hopkins University Press.

Lewis, J.E., DeGusta, D., Meyer, M.R., Monge, J.M., Mann, A.E., and Holloway, R.L. (2011). The Mismeasure of science: Stephen Jay Gould versus Samuel George Morton on Skulls and Bias. PLoS Biol 9(6): e1001071. doi: 10.1371/journal.pbio.1001071.

Nyborg, H. (2012). In Conversation with J. Philippe Rushton. *Personality and Individual Differences*, Special issue: The life history approach to human differences: J. Philippe Rushton in Memoriam (Part I in this book),

Nyborg H. (2003). The sociology of psychometric and biobehavioral sciences: A case study of destructive social reductionism and collective fraud in 20th century academia. In Nyborg H. (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 441-501). Amsterdam: Pergamon.

Nyborg, H. (2011). The greatest collective scientific fraud of the 20th Century: The demolition of differential psychology and eugenics. *Mankind Quarterly, 51,* 241-268.

Jensen, A.R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39,* 1-123.

Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human Nature.* New York: Viking Penguin.

Rushton, J.P. (1995). *Race, evolution, and behavior: A life-history perspective*. New Brunswick: Transaction.

Rushton, J.P. (1997). Race, intelligence, and the brain: The errors and omissions of the "revised" edition of S.J. Gould's *The Mismeasure of Man* (1996). *Personality and Individual Differences, 23*, 1, 169-180.

Rushton, J.P. (2010). Brain size as an explanation of national differences in IQ, longevity, and other life-history variables. *Personality and Individual Differences, 48*, 97-99.

Rushton, J.P., & Ankney, D.C. (2000). Size matters: a review and new analyses of racial differences in cranial capacity and intelligence that refute Kamin and Omari. *Personality and Individual Differences, 29*, 591-620.

Rushton, J.P. & Jensen, A.R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*, 235-294.

Segerstråle, U. (2000). *Defenders of the truth: The battle for science in the sociobiology debate and beyond*. Oxford, U.K.: Oxford University Press.

In conversation with J. Philippe Rushton

Helmuth Nyborg
Professor emeritus., Dr. Phil.
University of Aarhus, Denmark (1968-2007)

Helmuth Nyborg: Tell us about your background.

J. Philippe Rushton: I had one famous ancestor, Samuel Crompton (1753-1827) of Bolton, Lancashire, who invented the spinning mule which helped revolutionize the cotton-textile industry. It spun almost every type of yarn and was more versatile than the spinning jenny it replaced. Crompton lived in fear of the Luddites, bands of English workers (1811-1816) who went around the countryside smashing up new machines (and their inventors) lest they threaten jobs and the existing order of life. In the end, he was hailed as a benefactor and his threads and yarns became much sought after.

HN: Are all your roots in England?

JPR: Other than old Crompton, my ancestors were plain English working class folk, dissenters from the Established Church of England, and often anti-establishment in politics as well. The Rushton family name, along with their collaterals, the Ashcrofts and Cromptons, go back to the Saxon Chronicles (10^{th} century), perhaps entering England from Germany as early as the 4th century. My father's family had lived in Lancashire for generations as small farmers or local artisans. Crompton himself was the son of a tenant farmer near Bolton, which was the town my parents grew up in. My mother's family, the Adamsons, originated in Scotland, but the

family had lived in Lancashire for at least three generations. In World War I (1914-1918) my maternal grandfather served in the 10th Hussars, a Cavalry regiment, and was stationed in Northern France where he met my maternal grandmother, a Catholic farmer's daughter. So my mother contributed my middle name, spelt the French way.

HN: Tell us about your formative years.

JPR: I was born on December 3rd 1943 in the middle of World War II (1939-1945) in Bournemouth, a middle-class seaside resort in southern England. My father joined the Royal Air Force as ground crew, repairing battle-damaged Spitfires, and seeing action on airfields east of London during the 1940 Battle of Britain. My mother worked for the Fire Service in London; and an uncle served in the 8th Army under General Montgomery at the Battle of El Alamein. In 1943, my mother relocated to the family's pre-war house in Bournemouth and my father volunteered to work in a nearby armaments factory to be closer to her. I was the first result. My brother Peter soon followed in June 1945 as the war came to a close.

In 1945, a Labour Party (socialist) government was elected in Britain. It nationalized important industries and rationed the building supplies my father needed as a private contractor and disadvantaged him in competing with the massive state-run housing projects. With the future in Britain looking bleak for my small-business-oriented family, they decided to seek more fertile fields. In 1948 they emigrated to South Africa.

In 1952 we returned to the UK. I passed the 11+ examinations and began grammar school (high-schools offering advanced level courses). In 1956 my father landed a job he really wanted as a scenic artist and designer in the new Canadian Broadcasting Corporation (CBC) television stations in Toronto. My parents emigrated anew, moving the family to Canada.

HN: What drew you to psychology?

JPR: As a teenager I read Hans Eysenck's three popular paperbacks: *Uses and Abuses of Psychology* (1953), *Sense and Nonsense in Psychology* (1956), and *Fact and Fiction in Psychology* (1964). They brought a hard-nosed philosophy of science to real life topics like IQ testing, vocational interests, personality, politics, and psychotherapy. I used one of his questionnaires and mapped the political attitudes of my family and friends on his two axes of radical *versus* conservative, and tough-*versus* tender-minded. It was a revelation to find that people could be classified in such a straightforward manner. I turned out to be tough-minded, but neither conservative nor radical. At university, I was disappointed to find individual differences constituted only a tiny part of the syllabus, and were often considered "passé."

I returned to England to go to university and in 1970 earned a B.Sc. in psychology from Birkbeck College at the University of London with First Class Honors, and then in 1973 a Ph.D.in social psychology at the London School of Economics and Political Science. I was determined to get to grips with a substantial problem and make a contribution to knowledge. I chose "altruism" as my dissertation topic because it posed a fundamental problem for theories of human nature. "Why *do* people help others?" At the time I was enamored with social learning theory and thought if it explained altruism, it could explain other parts of human nature and have implications for improving the human condition.

I gained access to the local schools and examined generosity in 7- to 11-year- olds. Mine was an experimental study in which I varied whether children observed models behaving generously or selfishly. My supervisor, Hilde T. Himmelweit, Chair of the Department, herself an Hans Eysenck Ph.D. was best-known for her 1958 book, *Television and the child*. Although she would have preferred me to conduct surveys rather than "laboratory-type" experimental studies, she fully allowed me go my own way.

I spent 1973-74 at the University of Oxford on a post-doc with Jeffrey Gray, another Eysenck Ph.D, and continued my research on personality development in children. Having published four articles from my Ph.D. augured well for gaining the academic

post to which I aspired. I returned to Canada in 1974 and taught at York University (1974-1976) and the University of Toronto (1976-1977). I then moved to the University of Western Ontario and became a full professor in 1985. In 1989 I was made a Fellow of the John Simon Guggenheim Memorial Foundation and in 1992 I received a D.Sc. from the University of London (an earned degree).

HN: You continued to publish social learning studies of prosocial behavior while at York University and the University of Toronto, including how children learn to resist temptation. So when and how did you become a sociobiologist or 'evolutionary psychologist' to use the gentler and now accepted term?

JPR: Altruism has always been *the* central issue for me. However, in the first ten years of my research career (1970-1980) I studied it from a social learning perspective. When I began at Western, my first priority was to complete a book I had started at Oxford, *Altruism, socialization, and society* (1980), in which I describe the influence of the educational system, the mass media, and the family. I identified: empathy and internalized "norms of social responsibility" as primary motivations. So, I wasn't always a controversial race-researcher—or even an evolutionary psychologist.

Then, around 1979 I joined a group of colleagues from my department working our way through E. O. Wilson's (1975), *Sociobiology: The new synthesis*. Wilson defined altruism as the "central theoretical problem of sociobiology" (p. 3) and provided the answer (following Darwin) in family structure and kinship. Reading Wilson's tome with my colleagues (we met one evening a week at the Faculty Club) led me to the over-arching structure of evolutionary *r-K* life history theory. I adopted this perspective for my research from then on.

HN: But how did you develop that interest in sociobiology in the first place?

JPR: I underwent a "paradigm shift" by solving two anomalies from my social learning perspective with which I wrestled while writing the *Altruism* book. The first was whether altruism existed in animals. Wilson (1975) showed altruism in non-human animals was pervasive. For me, this implied an evolutionary and genetic basis in people. The evidence that prosocial parents produce prosocial offspring (and abusive parents, abusive offspring), might just as easily be explained genetically as culturally. When reviewing the literature on family influences, I was struck by how few studies controlled for genetic influences. It wasn't until the 1980s that genetic designs began to be incorporated more routinely into social developmental research.

The other anomaly was whether altruism was an enduring trait of personality. If people were as consistent in their prosocial behavior as they were in intelligence and temperament then it would be correct to describe them as having traits of character. Though it might seem commonsense to ascribe traits to individuals, a major debate during the 1960s and 1970s arose over the "consistency versus specificity" of behavior. As a social learning theorist I typically sided with the specificity position, although I always accepted the three super-traits of Intelligence, Extraversion, and Anxiety. It was a major intellectual reversal (indeed an epiphany) when I concluded that people did indeed have enduring traits of character. To examine them one had to use the *principle of aggregation* and form composite measures (Rushton, Brainerd & Presley, 1983).

Solving these two anomalies to my own satisfaction—altruism in animals and traits in people—led me from social learning to sociobiology.

HN: So, to "retool," you took a semester away from your home university and spent January to June 1981 as a Visiting Scholar at the University of California, Berkeley?

JPR: Yes. I had been invited to the Institute of Human Development at Berkeley by fellow altruism researcher Paul

Mussen. Specifically, I hoped to examine how stable individual differences in altruism were over the life-span. However, I found that although many of the Institute's members had international reputations documenting the early emergence of personality traits in children and their power to predict adjustment, no one was interested in discussing genetic causation. At Berkeley, discussion of behavior genetics was a hop, skip, and a jump away from Arthur Jensen's politically incorrect research on Black-White IQ differences.

HN: One question I am asked about you is why you picked such a controversial topic to study as race differences, and then stayed with it for so long? Is it that you enjoy being in the limelight? Marvin Zuckerman once called you an "intellectual sensation seeker."

JPR: Well I do enjoy intellectual excitement, but I never expected my work on generosity in children to lead to a firestorm over race differences. In the end I felt it couldn't be avoided if there was to be a full evolutionary analysis of human life histories.

HN: But surely it could have been avoided? Most behavior geneticists and evolutionary psychologists study human behavior without getting embroiled in race differences in brain size and intelligence? Why didn't you steer clear of the minefields?

JPR: Perhaps it might have gone differently if I hadn't been as strongly influenced by Arthur Jensen. Since he occupied an office in the School of Education, one floor up from my office in the psychology department, I decided to visit him and we hit it off. I had been interested in his work on race and intelligence since graduate school, though I remained agnostic as to any genetic basis. Jensen was highly informative, sketching out his views and providing detailed answers to my questions along with copies of his reprints. Under his tutelage, I learned about general intelligence (Spearman's *g*), behavior genetics, and race differences.

Jensen's (1969, 1973) work went beyond IQ. He documented that, on average, Black babies are born a week earlier than White babies yet are more mature on measures of amniotic fluid in the placenta, bone development by X-rays, and muscular strength and motor co-ordination. East Asian infants are less precocious and motorically reactive than Europeans. He also described differences in two-egg twinning rates (16, 8, and 4 per 1,000 live births for Africans, Europeans, and East Asians, respectively). In a long footnote, Jensen (1973) suggested that "the three racial groups lie on a developmental continuum on which the Caucasian group is more or less intermediate (p. 289)."

Jensen's note struck a responsive chord for I had been reading about *scala naturae* and behavioral scaling in the sociobiological literature and wondered about their application to human differences. Wilson (1975) explained the origins of parental care, which had enormously increased in complexity over evolutionary time (Figure 1). He described two ends of a reproductive continuum. At one end, a "fast" life history (the *r*-strategy), eggs and sperm are produced and simply discharged into the water (for example, in oysters). Further to the opposite end, a "slow" life history (the *K*-strategy), an egg is not only laid in the ground but pollen and honey provided for future needs (as with wasps). Other steps in the *K* direction include bringing food and caring for the offspring. In mammals, the combined demands of gestation, delivery, production of milk, and protecting and physically caring for the young reach a peak.

K-strategists provide a lot of parental care. They have complex social systems and work together in getting food and shelter. *K*-strategists have more developed nervous systems and bigger brains but produce fewer eggs and sperm. The bigger an animal's brain, the longer it takes to reach sexual maturity and the fewer offspring it produces. Number of offspring, time between births, parental care, infant mortality, speed of maturity, life span, even social organization and altruism all fit together like pieces of a puzzle.

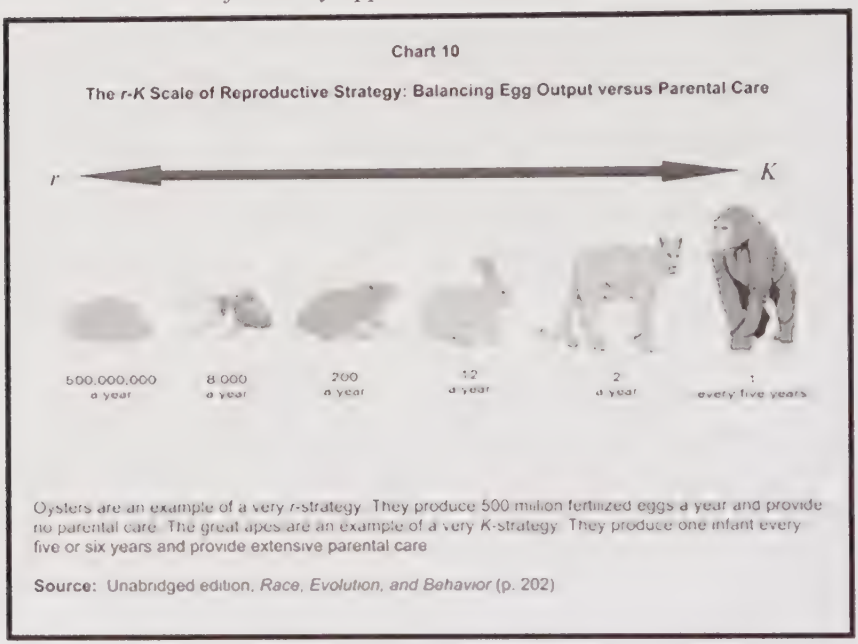


Figure 1. The r-K scale of reproductive strategy.

I'd also read an article on the "Origin of Man" by C. Owen Lovejoy (*Science*, January 31, 1981) which documented the trend in hominids toward larger brains, prolonged gestation, single births, a longer period of infant dependency, successively greater periods between pregnancies, greater predator recognition and territorial defense, and prolonged life-spans. Lovejoy attributed this pattern to an increasingly *K*-type demographic strategy. With each step in the scale, populations require a greater proportion of their reproductive energy to be devoted to sub-adult care, with greater investment in the survival of fewer offspring. I couldn't help but wonder whether the well-documented racial differences in family structure might have their roots in just such an origin?

I reviewed the international literature on differences among the three major races on 60 different variables (Table 1). Some were well-established, such as family structure, crime, and educational

achievement. East Asian Americans, often labeled a "model minority," averaged more tightly knit and complexly organized families than European Americans. Other data had only recently become available such as Richard Lynn's (1977) finding that East Asians averaged a higher IQ than Europeans. Some data were sketchy. Were there *really* race differences in sexual behavior—as the stereotypes implied? I later found race differences did exist in age of sexual maturation (Blacks earlier than Whites and Whites earlier than East Asians), frequency of intercourse and masturbation, as well as in sexual anatomy.

Table 1. Average differences between Africans, Europeans and East Asians.

	Africans	**Europeans**	**East Asians**
Brain size			
Mean across methods (cm^3)	1,267	1,347	1,364
Autopsy data (cm^3 equivalents)	1,223	1,356	1,351
Endocranial volume (cm^3)	1,268	1,362	1,415
External head measures (cm^3)	1,294	1,329	1,356
Cortical neurons (billions)	13,185	13,665	13,767
Intelligence			
IQ scores	70-85	100	105
Decision times	Slower	Intermediate	Faster
Cultural achievements	Lower	Higher	Higher
Muscular-skeletal traits			
Muscle attachment sites on crania	Largest	Intermediate	Smallest
Postorbital constriction and temporalis fossae (indentations in skull for jaw muscles)	Largest	Intermediate	Smallest
Facial prognathism (forward jutting jaw)	Most	Intermediate	Least
Number of teeth	32	30-32	28-30
Size of molars	Largest	Intermediate	Smallest
Bi-condylar breadth of mandible (widening of upper back-of-jaw for attachment to wider skull)	Least	Intermediate	Largest
Mass of nuchal muscles	Largest	Intermediate	Smallest
Femoral head size (where thighbone exits pelvis)	Smallest	Intermediate	Largest

I considered data on brain size to be especially important. In 1981 there was an almost empty cell in Table 1 for East Asians, although *r-K* theory predicted they should average higher. Then data showing an East Asian advantage were provided by Beals, Smith, and Dodd (1984). In the 1990s I corroborated this advantage by calculating cranial capacities for international samples of air-force pilots collated by NASA, a stratified random sample of 6,325 U.S. Army personnel, thousands of 25- to 45- year olds grouped into 40 regional samples by the International Labour Office, and 35,000 children followed from birth to 7 years by the U.S. National Collaborative Perinatal Project (see Rushton, 1995; Rushton & Ankney, 2009).

The literature on hominid evolution showed that as brain size increased from chimpanzees (380 cm^3) to *Australopithecus* (450 cm^3) to *H. habilis* (650 cm^3) to *H. erectus* (1,000 cm^3) to *Homo sapiens* (1,400 cm^3), it was accompanied by changes from a robust to a more gracile form across 74 musculo-skeletal traits, both cranial and postcranial (r = .90 across the populations). Subsequently, I examined 37 of the 74 traits representing most of the skeleton to determine whether the trend applied across the three sub-species of *H. sapiens* and found that it did (Table 1; Rushton & Rushton, 2003). The initial 74 traits across five species were taken from evolutionary anatomy textbooks; the 37 on race differences from forensic anthropology textbooks. Across the three populations, the 'ecological correlation' between brain size and the 37 traits averaged a remarkable r = .80. If the races did not differ in brain size, this correlation could not have been found.

An unexpected finding was that males averaged a larger brain than females even though they did not have a higher IQ. Richard Lynn (1999) dubbed this the "Ankney-Rushton anomaly," after my zoology colleague Dave Ankney discovered the sex difference, even after controlling for body size (see Rushton & Ankney, 2009). Lynn solved the paradox by showing that men averaged 5 IQ points higher than women. He suggested the sex difference had gone unnoticed because females mature faster than males thereby masking differences prior to late adolescence when

most IQ data is collected. Lynn's finding of a higher male IQ after adolescence to go along with the larger brain has been corroborated by others, including Doug Jackson and myself in an analysis of 100,000 SAT takers (Jackson & Rushton, 2006).

HN: It's still not clear why you made race the center of your research. You titled your 1995 book, Race, Evolution, and Behavior.

JPR: The question of why the *three-way pattern* of racial differences had come about (Table 1) was far too intriguing to walk away from. It offered more opportunity for finding and testing alternative theories than any single dimension. Only a gene-based evolutionary theory could explain the totality of this pattern.

HN: So was the application of r-K life history to human differences unique to you? It must have taken a lot of hard thinking to work it all through. How did you make the leap?

JPR: Remarkably, there was no need to think it through. Instead there were a series of ever stronger "clang associations," mostly in 1980 and early 1981, until a light went on in my head and I thought "Hmm, that seems to fit; I wonder how I might prove (or disprove) it." The answer was again immediate–examine whether other *r-K* traits lined up in the same direction. It didn't take more than ten minutes to list enough items to see the approach was viable (Table 1).

HN: Is the General Factor of Personality (GFP) part of the r-K life-history approach? If so, how?

JPR: Yes. In my first publication on *r-K* life history theory, I proposed that "one basic dimension—*K*—underlies much of the field of personality" (Rushton, 1985, p. 445). Subsequent research confirmed the interrelatedness of the life history traits identified in the 1985 paper such as altruism, intelligence, sexuality, and

delinquency. But, it probably was Sir Francis Galton (1887) who first described a GFP, just as he did the general factor of intelligence (see Rushton & Irwing, 2011).

The GFP as such, however, only became known in 2007 after hierarchical factor analyses were carried out on existing personality inventories assessing both normal and personality disorder scales (Rushton, Bons, & Hur, 2008; see Rushton & Irwing, 2011, for review). Cross-national twin studies find 50% of the variance in the GFP is due to genetic influence and 50% to non-shared environmental influence. A South Korean twin study found the GFP emerged by 2- to 3-years of age (Rushton et al., 2008). As with cognitive ability, an integration of broad and narrow traits can be achieved through combining them hierarchically; with the GFP occupying the apex in the same way that *g* occupies the apex in the organization of cognitive abilities.

High scores on the GFP indicate a "good" personality; low scores a "difficult" personality (someone who is hard to get along with). Individuals high on the GFP are altruistic, agreeable, relaxed, conscientious, sociable, and open, with high levels of well-being and self-esteem. Because the GFP defines clear positive and negative poles, it provides potential for understanding the socially "advantaged" (those with high levels of emotional intelligence) as well as the socially "challenged" (those more likely to suffer a personality disorder). The GFP can be viewed as a dimension of social effectiveness. This follows Darwin's proposal that natural selection acted directionally to endow people with more cooperative and less contentious personalities than their archaic ancestors or nearest living relatives, the chimpanzees. Individuals high on the GFP probably left more progeny since we prefer mates, fellow workers, and leaders who are altruistic, conscientious, and emotionally stable. Those able to cooperate in groups were also more likely to achieve goals and even win wars.

HN: You have also made contributions to altruism, personality, and creativity. For example, you formulated Genetic Similarity Theory by which you mean that people innately prefer those genetically

similar to themselves. Can you expand on this and your views on assortative mating?

JPR: I began to work on what I called Genetic Similarity Theory (GST) in late 1981 after returning to Canada from Berkeley. There I had been struck by how much people associated with those similar to themselves and the amount of 'ethnic nepotism' engaged in by political activists. Seeking an evolutionary explanation, I extended William Hamilton's (1964) inclusive fitness solution to the problem of altruism. It stated that organisms act to benefit kin depending on how closely related they are. However, in human beings, altruism goes well beyond kin (that is, 'blood relatives'). People tend to associate with, befriend, marry, and help others who, although not kin, are genetically similar (see Rushton, 2009, for review).

The evidence I amassed over the next 25 years in favor of GST included finding that: spouses and close friends assort on blood groups and that blood group similarity predicts fertility; twin and adoption studies show that genes, not upbringing, cause people to assort positively; phenotype matching studies show that assortment is more pronounced on the more heritable items within sets of physical and psychological traits; bereavement studies find that grief is greater following the death of a more similar co-twin or child; and studies of face perception find that people prefer and trust those who look like them.

Using Jensen's (1998) Method of Correlated Vectors, my studies found that social assortment was more pronounced on the more heritable components of traits. I found that attitude and personality items from 300 pairs of identical (MZ) and fraternal (DZ) twins, their spouses, and their best friends revealed the spouses and friends were more similar to each other than to first cousins and almost as similar as to brothers and sisters, a degree of similarity not previously recognized (Rushton & Bons, 2005). Further, the twins' preference for similarity in social partners was 34%, heritable, with 12% of the variance being due to family upbringing, and 54% due to chance factors. For unconscious

genetic reasons we seek out our 'own kind' - extraverts with extraverts; traditionalists with traditionalists. The reason is that by liking, becoming friends, aiding, and mating with those who are genetically similar, we help ensure that our own segment of the gene pool is maintained and transmitted to future generations.

HN: In the 2005 issue of Nations and Nationalism (an academic journal published by the London School of Economics) you described the implications of Genetic Similarity Theory for ethnic relations. Please explain.

JPR: The pull of genetic similarity does not stop at family and friends. Group members move into ethnic neighborhoods and join together in clubs, societies, and religious organizations. DNA analyses demonstrate a startlingly obvious result. Individuals from the *same* ethnic group share many more genes with their co-ethnics than they do with those of other ethnic groups. Based on their DNA, individuals randomly chosen from the same ethnic group are as much like their ethnic group as they are their extended families; they are like half siblings compared to individuals randomly chosen from the rest of humanity. Two random English people are equivalent to 3/8 cousins compared with two random people from the Near East; 1/2 cousins compared to two random people from India; half-sibs by comparison with people from China; and like full-sibs by comparison with people from sub-Saharan Africa. Rather than being a poor relation of family nepotism, ethnic nepotism is an extension of it. The evolutionary reason why altruism follows lines of genetic similarity is to replicate genes more effectively. Xenophobia may represent the dark side of the evolutionary coin. From an evolutionary perspective it can be rational for an individual to sacrifice for the group. In extreme cases such as war, it can be more important than helping immediate kin (See Salter & Harpending, this issue).

HN: You also published twin studies using the University of London Twin Register in the U.K. The first of these was in collaboration

with Hans Eysenck, with whom you spent a sabbatical in 1982-83 (Rushton, Fulker, Neale, Nias, & Eysenck, 1986). You found that individual differences in altruism, empathy, nurturance, and aggression were 50% heritable. In 1996 you published a paper showing the heritability of violent crime. In 2004, the Royal Society published your paper on the links between genes and altruism showing there were genetic reasons for human kindness (Rushton, 2004). Can you summarize your views?

JPR: I was surprised that both altruism and anti-social behavior had such high heritability because I thought they would be subject to much more socialization pressure than neutral traits. The 2004 study analyzed social attitudes in 300 pairs of twins and found people's heritable tendency to act in a socially responsible manner, help others and disapprove of cheaters, covered a broader array of situations than previously examined. I concluded that humans are hardwired with a "goodness gene" that forms a deep part of human nature. Society is not built to fall apart at the seams. A group of strangers on a desert island would soon band together and work toward a common good.

HN: In other research you found support for Eysenck's theory that creativity is linked to Psychoticism. Does this contradict the tendency for good traits to go together, as in the General Factor of Personality?

JPR: Creativity remains a conundrum for me and I still don't really know what to make of it in terms of the GFP. Jensen (1998) thought the Big Five traits would predict productivity, especially Conscientiousness and Openness, but also Agreeableness, Sociability, and Emotional Stability when getting on well with people mattered. He also argued that much "zeal and industry" were necessary for very high creativity. The ordinary term "motivation," he pointed out, seemed too intentional and self-willed to fit the behavior of individuals whose biographies showed them to be obsessed by the subject of their work. Jensen suggested that this

obsessive-compulsive mental activity in a particular sphere was beyond conscious control. Instead, he likened the "mental energy" of a genius to the kind of cortical arousal seen under the influence of stimulant drugs.

Although Jensen's traits can be fitted into the GFP, there is much evidence that, in addition to being obsessive-compulsive, highly creative individuals have a greater percentage of other psychological disorders. Eysenck's theory is that Psychoticism inclines people to generate quirky ideas, which predispose them on the one hand, to a personality disorder, and on the other, to greater creativity. Eysenck further postulated that *intelligence* enhanced creativity primarily through the *speed* with which new associations were formed, while *psychoticism* did so by expanding the *width* of the associations it could bring together, which he termed over-inclusive thinking. Thus, intelligence and psychoticism act as independent contributors to creativity.

In the 1970s while at York University I became interested in *scientometrics* and using citation counts to measure eminence. Part of my motivation was to see how I was doing personally, compared to colleagues, and how far I would have to go to excel. These are hardly the noblest of motives, but in defense I can say I was also very interested in scientific biography and what kind of people made great discoveries.

My first study in 1977, with Norman Endler and Roddy Roediger in the *American Psychologist*, confirmed what others had found, that very few scientists are responsible for producing the great majority of creative works. 52% of our sample did not publish a single article in any journal in 1975. The picture was similar for citations—53% of psychologists had 5 or fewer citations, while only 25% had 15 or more, and only 1% had 100 or more.
Subsequently, using both self- and other-ratings made by university professors (*N*s = 52, 69) we found several personality traits consistently correlated with publication and citation counts: ambition, endurance, seeking definiteness, dominance, showing leadership, aggressiveness, intelligence, independence, non-meekness, and non-supportiveness. I asked Eysenck to estimate the

loading of each of the 29 traits on his Psychoticism (*P*) scale, which he argued predicted schizophrenia and other psychotic disorders, and which he identified with traits such as high in aggressiveness, coldness, egocentricity, impersonality, impulsivity, and tough-mindedness. I used Eysenck's estimates to calculate an overall *P* score for each professor and correlated this with a composite of publications and citations. For the study of 52 professors, the correlation between *P* and the composite was $r = .26$ ($p < .05$; for intelligence it was $r = .40$, $p < .01$). For the study of 69 professors, the correlation was $r = .43$ ($p < .01$; for intelligence it was $r = .05$). I carried another study with 200 university students using the *P* scale from the Eysenck Personality Questionnaire along with a measure of general intelligence and a creativity test based on Divergent Thinking (Rushton, 1990). Both *P* and IQ correlated positively with the total number of ideas generated on the measure of divergent thinking ($r = 17$, $p < .05$; $r = .24$, $p < .05$, respectively). In all three studies, Psychoticism was unrelated to Intelligence and thus made an independent contribution to creativity. This evidence, too, confirmed Eysenck's hypothesis.

HN: Any last thoughts?

JPR: Charles Darwin and E. O. Wilson were correct. Human social behavior is best understood as part of a life history—a suite of traits genetically organized to meet the trials of life—survival, growth, and reproduction.

References

Beals, K. L., Smith, C. L., & Dodd, S. M. (1984). Brain size, cranial morphology, climate, and time machines. *Current Anthropology*, *25*, 301–330.
Eysenck, H. J. (1953). *Uses and abuses of psychology*. London: Penguin.
Eysenck, H. J. (1956). *Sense and nonsense in psychology*. London: Penguin.
Eysenck, H. J. (1964). *Fact and fiction in psychology*. London: Penguin.
Eysenck, H. J. (1995). *Genius: The natural history of creativity*. Cambridge: Cambridge University Press.
Galton, F. (1887). Good and bad temper in English families. *Fortnightly Review*, *42*, 21-30.
Hamilton, W. D. (1964). The genetical evolution of social behavior: I and II. *Journal of Theoretical Biology*, *7*, 1-52.
Jackson, D. N., & Rushton, J. P. (2006). Males have greater *g*: Sex differences in general mental ability from 100,000 17- to 18-year-olds on the Scholastic Assessment Test. *Intelligence*, *34*, 479-486.
Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*, 1-123.
Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.
Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger.
Lovejoy, C. O. (1981). The origin of man. *Science*, *211*, 341-350.
Lynn, R. (1977). The intelligence of the Japanese. *Bulletin of the British Psychological Society*, *30*, 69–72.
Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences*, *17*, 257-271.
Rushton, J. P. (1980). *Altruism, socialization, and society*. Englewood Cliffs, NJ: Prentice Hall.

Rushton, J. P. (1985). Differential *K* Theory: The sociobiology of individual and group differences. *Personality and Individual Differences, 6*, 441-452.
Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective.* New Brunswick, NJ: Transaction.
Rushton, J. P. (2004). Genetic and environmental contributions to prosocial attitudes: A twin study of social responsibility. *Proceedings of the Royal Society of London, B, Biological Sciences*, *271*, 2583-2585.
Rushton, J. P. (2009). Inclusive fitness in human relationships. *Biological Journal of the Linnean Society*, *96*, 8-12.
Rushton, J. P., & Ankney, C. D. (2009). Whole-brain size and general mental ability: A review. *International Journal of Neuroscience*, *119*, 691-731.
Rushton, J. P., & Bons, T. A. (2005). Mate choice and friendship in twins: Evidence for genetic similarity. *Psychological Science, 16*, 555-559.
Rushton, J. P., Bons, T. A., & Hur, Y-M. (2008). The genetics and evolution of a general factor of personality. *Journal of Research in Personality*, *42*, 1173-85. (Corrigendum, 2009, *Journal of Research in Personality*, *43*, 532.)
Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94,* 18-38.
Rushton, J. P., Fulker, D. W., Neale, M. C., Nias, D. K. B., & Eysenck, H. J. (1986). Altruism and aggression: The heritability of individual differences. *Journal of Personality and Social Psychology*, *50*, 1192-1198.
Rushton, J. P., & Irwing, P. (2011). The general factor of personality: Normal and abnormal. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell Handbook of Individual Differences* (pp. 132-161). London: Wiley-Blackwell.
Rushton, J. P., & Rushton, E. W. (2003). Brain size, IQ, and racial-group differences: Evidence from musculoskeletal traits. *Intelligence, 31,* 139-155.

Wilson, E. O. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Harvard University Press.

Part 2: The Book

Introduction

This special issue –*The life history approach to human differences: J. Philippe Rushton in Memoriam* - was originally intended as a collegial tribute to Professor J. Philippe Rushton, University of Western Ontario, Canada, to be handed over to him on a special occasion where colleagues would personally testify to his many unique contributions to science. This was not going to be so. It is with deep regret we received word of Phil's untimely death on 2. October 2012.

We knew that he struggled against Addisons disease, which increasingly compromised the immune system and made it too risky for him to attend large collegial meetings, but we at least hoped he would be able to receive our tribute in the company of a small group of admirers. Instead we must turn this special issue into a memorial collection of works.

The issue consists of 15 papers, divided into four parts: **The Man**; **Introduction**; **Part III** with papers on Rushton's contribution to the study of differences in intelligence; and, **Part IV** containing papers discussing Phil's many contributions to the analysis of **Personality and Development**, including altruism, personality structure, life history theory, genetic similarity theory, nepotism, cross-national mesures of penis length, migratory selection, global variation in differential-K, and dysgenic fertility.

The first part on **The Man** presents an Obituary for Phil, as he regrettably died 2 October 2012, and thus never got a chance to laid eyes on this collection of papers in his honour. Luckily, we managed to chronicle an in depth interview with him shortly before his untimely death. It follows after the Obituary, and provides a glimpse into his personal and professional background and asking about his own reflections on the many areas of interests he

happened to pursue successfully during his too short and often troubled career.

In the third part on intelligence, Arthur Jensen first outlines Rushton's many innovative ways of confirming the "Spearman-Jensen hypothesis", i.e. that Black-White difference in general intelligence is *g*-loaded. Jensen describes how Rushton confirmed this hypothesis – often called a *Jensen Effect when showing it by means of the correlated vector approach* - in various other samples, how he documented significant average group differences in a suite of essential biological and social traits, and how Rushton explained the overall scheme in terms of life history theory.

Linda Gottfredson explains Rushton's pioneering life-history research, which he applied to the human species more systematically than anyone else, and then takes task with some of the vitriolic critique of it. She outlines the derogative nature of "mob science", exposes its enforced collective ignorance, and explains how the systematic attacks are used to quarantine not only Rushton's human biodiversity research, but Rushton the scientist as well.

The following paper by Jan te Nijenhuis focus on the Flynn-effect (secular IQ gains) and probes the Spearman-Jensen hypothesis, which Rushton defends, in several datasets, and finds a small negative average correlation between score gains and *g*-loadings. Nijenhuis takes this to mean that the Flynn- effect and group differences in intelligence most likely have different causes, but also underlines the need for a psychometric meta-analysis of all available studies.

The last paper in this section is by Heiner Rindermann. He checks the validity of past studies of intelligence in Africa by integrating various cognitive ability measures, and comes to the conclusion that the best estimate of African average IQ is 75. He also discusses environmental versus evolutionary factors relevant for intelligence, ventures into how to enhance African cognitive development, and ends by suggesting that the Black-White gap of about 1 SD will narrow due to modernization and overcoming current environmental obstacles.

The first paper in Part IV on **Personality and Development** by Paul Irwing from Manchester University, UK, concentrates on the observation that a general factor of personality (GFP) is extractable from most data on personality. Such extraction has been criticized for various reasons, and Irwing first discusses each of the objections and then provides new evidence and arguments in favor of the existence of a GFP.

Then Donald Templer from Alient International University in California evaluates Rushton's role as a theoretician and amalgamator of individual and group differences in personality with Differential K-, and Life-History theory. In addition, Templer proposes his own two-pillar theory, which is based on psychometric *g* and the general factor of personality, GFP. The theory is heavily inspired by Rushtonian thinking, but emphasizes that where biology determines one's life course, life events produce deviations.

Yoon-Mi Hur from Mokpo National University in South Korea treats one of Rushtons persistent interests, altruism, and exemplifies how Rushton moved from a social learning perspective to a sociobiological view, consistent with 50% inheritance. She also details Rushton's view on the GFP as a dimension of social effectiveness, which dovetails nicely with Darwin's view that natural selection gradually endowed humans with more cooperative and less contentious personalities.

Figueredo, de Baca, and Woodley from University of Arizona evaluate Rushton's application of life history theory when he explained the striking co-variation among human traits in terms of the latent structure of life history, covitality, and personality. They conclude that Rushton pioneered in the study of individual differences by proposing a theoretically meaningful and empirically useful new paradigm.

Frank Salter from University of Sydney, Australia, and Henry Harpending from University of Utah, USA, outline and evaluate Rushton's genetic similarity theory in terms of ethnic nepotism and other life history aspects. They find that his theory is rather promising for further research in evolutionary social science

because it "... unifies evolutionary and behavioral mechanisms in a single theory".

Richard Lynn from University of Ulster in Northern Ireland surely moves into a scientific mine-field by presenting empirical data on race differences in penis size and considering them in terms of Rushton's *r-K* life history theory. It is curious, at least to me, that particular measures of certain human body parts can raise such strong emotions, as Rushton ran into when he in 2000 first reported race differences in penis length and diameter. Now, Lynn provides new data on racially differently disaggregated measures of penile size and is able to generally confirm predictions based on Rushtonian theory. This is taken, in evolutionary life history terms, to mean that differences in penis size and IQ have evolved independently in response to climatic challenges, population density, and propensity for infidelity.

Helmuth Nyborg from Aarhus University (1968-2007), Denmark, proposes an IQ/T Geo-Climatic theory to explain how primordial migration out of Africa selected increasingly for higher IQ and lower aggression via lower testosterone as the climate got colder. Where Rushton explains similar phenomena in terms of life history theory, Nyborg envisions this migratory transaction in terms of organisms having to deal with changing geo-climatic selective pressures, but are restricted by a fixed intra-systematic energy budget: The "expensive" development of a large intelligent brain in the cold and costly establishment of associated *g*-Nexus traits, therefore had to be compensated for by less development of T–Nexus traits through lowering the plasma testosterone level. Test of the predictions of the General Trait Covariance model supports Rushton's three-racial ranking of covariant trait patterns. Nyborg argues that the recent massive South-to-North immigration reverses this primordial evolutionary progression and leads Western civilization into decay.

Gerhard Meisenberg and Michael A. Woodley from Ross University School of Medicine in Dominica operationalize Rushton's differential-*K* spectrum at the level of countries in their paper. This national *K* factor correlates 0.877 with national IQs.

They find that *K* is predicted by IQ, log(*GDP*), and skin color, and further that general intelligence, *g*, is predicted by skin color, log(*GDP*), cranial capacity, and evolutionary novelty. Meisenberg and Woodley conclude that evolutionary conditions are plausible contributors to the current worldwide distribution of K and IQ.

The last paper by Michael A. Woodley and Gerhard Meisenberg attempts to determine if dysgenic fertility relates to the Jensen-Effect, that is, whether IQ subtests with high *g*-loadings relate to higher dysgenic fertility, and subtests with low loadings less so. Applying Jensen's method of correlated vectors to various race- and sex-samples, they find a significant Jensen effect in five out of seven samples. They conclude that dysgenesis occurs on the "genetic *g*" – and not on the Flynn effect (which is "hollow" with respect to *g*). Woodley and Meisenberg conclude that genotypic IQ or "genetic *g*" has declined significantly over time.

Part 3: Intelligence

Chapter 1

Rushton's Contributions to the Study of Mental Ability

Arthur R. Jensen [1, 2]
[1] *University of California, Berkeley.*
[2] *This is regrettably the last paper from Arthur Jensen's hand. He died only 14 days apart from Phil Rushton. References to some Obituaries, and to a collegial tribute to Art's monumental works over the years, are provided at the end of this chapter.*

Abstract
This essay describes Rushton's contribution to examining the nexus of intelligence, race, and genetics, specifically what I termed "Spearman's hypothesis." It states that Black-White differences are "most marked in just those [tests] which are known to be saturated with *g*." I (Jensen) had confirmed this hypothesis using large data sets in the 1970s and 1980s and also found that Black-White differences were most marked on the more heritable rather than the more cultural subtests. Rushton confirmed and extended these findings in many highly innovative ways and demonstrated Spearman's hypothesis applied among samples of Gypsy Roma in Serbia, and East Asian, European, South Asian, Colored and Black samples in South Africa. He has not only documented group differences in brain size, intelligence, life span, family structure, infant mortality, developmental precocity, personality, and temperament, and rates of two egg twinning, and crime among East Asians, Europeans, and Africans, but also provided a life history theory that explains them.

1. Rushton's Contributions to the Study of Mental Ability

I first became acquainted with Phil Rushton in 1981 while he was on sabbatical at the University of California at Berkeley, where I had been a professor of education since 1964. When he came to see me in my office at the Institute of Human Learning, we hit it off immediately. Perhaps because I had done a post-doc with Hans Eysenck at the UK's Institute of Psychiatry in the University of London and Rushton had done his B.Sc. and Ph.D. at the University of London (although not with Eysenck), we shared the same approach to the study of individual differences. Often called the "London School," the trajectory went back through Hans Eysenck (1916-1997), Sir Cyril Burt (1883-1971), Charles Spearman (1863-1945) and Karl Pearson (1857-1936), all the way to Charles Darwin's (1809-1882) cousin, Sir Francis Galton (1822-1911), the founder of differential psychology (see Miele, 2002; Nyborg, 2003). These were the founding fathers of behavior genetics, statistical methods, and of my own present focus, mental chronometry [reaction time] (Jensen, 2006, 2011).

I was impressed not only by Rushton's intellect, but his intellectual boldness as well. There seemed to be no question in behavioral science, however vexing or controversial, that he was unwilling to ask me about and later tackle. And he did so with the three characteristics that our mutual inspiration, Galton, identified as the defining characteristics of genius—ability, zeal, and industry. Rushton would display these in full measure in the years that followed. However, he has done far more than simply accumulate inductive data. Accumulate he has, but he has also used the hypothetico-deductive method in making novel predictions to test hypotheses.

2. Spearman's hypothesis and the nature of *g*

I had demonstrated in my research of the 1970s that mean Black-White differences in IQ were more pronounced on the more heritable, less cultural subtests. For example, Jensen (1973) cited a

study by Nichols (1972) which found a correlation of $r = .67$ ($p < .05$) between the heritabilities of 13 tests estimated from twins and the magnitude of the Black-White differences on the same tests. I further demonstrated an inverse relation of $r = -.70$ ($p < .01$) between the environmentality (the converse of heritability, that is, the percentage of variance that can be attributed to non-genetic factors) for 16 tests estimated from differences between siblings and the mean White–Black differences (Jensen, 1973).

Jensen (1998) also tested Spearman's hypothesis using 12 reaction time (RT) measures from 820 9- to 12-year-olds in California. Reaction time is one of the simplest and most culture free measures. Many RT tasks are so easy that 9- to 12-year-old children can perform them in less than one second. Yet even on these very simple tests, children with higher IQ scores do so faster than children with lower scores, and East Asian 9- to 12-year-olds are, on average, faster than Whites who are, again on average, faster than Blacks. Moreover, the differences between Blacks, Whites, and East Asians in average RTs are largely on the *g* factor, with the correlations between the *g* loadings and the mean group differences ranging from .70 to .81. Since school children are not trained on reaction time tasks, as they are on paper-and-pencil tests, the advantage of those with higher IQ scores on RT tasks is unlikely to arise from practice, familiarity, education, or instruction. Moreover, although East Asians averaged faster (cognitive) *decision* times than Whites or Blacks, Blacks averaged faster (motoric) *movement* times than Whites or East Asians, thereby arguing against the hypothesis that the differences on these tests reflect any difference in motivation or attention. These results bear out Spearman's hypothesis even more strongly than do those from conventional psychometric tests. The hypothesis that the group difference on these tests reflects a difference in motivation is also disconfirmed by the fact East Asians typically averaged higher than Whites on the *g* factor extracted from their (faster) reaction times.

While carrying out these studies I developed what I termed the *method of correlated vectors* (MCV). It tests whether there is an association between a column of quantified elements (such as

the *g* loadings for a set of tests) and a parallel column of any independently derived scores (such as pass rate differences between groups). This technique produced a large number of results which made it clear that Spearman's *g* is the "active ingredient" in IQ scores. A test's *g* loading is embedded to a greater or lesser extent in every question on a mental test and is the best predictor, not just of that test's correlation with scholastic and work-place performance, but a number of *biological* measures including: heritability coefficients determined from twin studies, inbreeding depression scores calculated in children of cousin marriages, brain evoked potentials, brain pH levels, brain glucose metabolism, as well as nerve conduction velocity, and reaction time measures. These correlations argue strongly for the heritable and biological as opposed to the mere statistical reality of *g* (Jensen, 1998).

Rushton (1998) was kind enough to suggest the term "Jensen Effect" be used whenever a significant correlation could be demonstrated between *g* factor loadings and any other variable. Such "Jensen Effects" are not omnipresent and their absence can be as informative as their presence. For example, the Flynn Effect (the secular rise in IQ) is *not* a Jensen Effect because it does not occur on *g* (Rushton & Jensen, 2010).

Rushton (1989) carried out a particularly innovative and revealing study using inbreeding depression scores on subtests of the Wechsler Intelligence Scale for Children (WISC) in place of the usual *g* loadings and found remarkable generalization across populations. The inbreeding depression scores had been calculated by geneticists for each of 11 subtests from a study of Japanese twins. Using the method of correlated vectors, Rushton found these predicted the magnitude of the Black-White differences on the same subtests in the US ($r = .48$; $p < .05$). Inbreeding depression is an established genetic phenomenon that occurs when people who are genetically related have children together thereby producing in their offspring, on average, a lower score on IQ (or any heritable measure) than would otherwise have been the case. This occurs because of any combination of deleterious genes from both sides of the family. Usually kept recessive, these now combine to produce

a dominant allele and the consequent loss of about 5 IQ points in the offspring of cousins. There is no known non-genetic explanation for the phenomenon of inbreeding depression or why inbreeding depression scores in Japan should predict the magnitude of Black-White differences in the US.

Rushton also demonstrated Jensen Effects in Southern Africa for differences among South Asians (East Indians) and "Coloreds" (the term used to refer to the mixed-race population of South Africa) as well as for Whites and Blacks. In one study, Rushton (2001) reanalyzed data on 10 subtests of the WISC–R published by Skuy, Schutte, Fridjhon, and O'Carroll (2001) on 154 Black South African high school students. Rushton found the African–White differences were mainly on *g*. Rushton and Jensen (2003) compared data on the WISC–R from 204 African 12- to 14-year-olds from Zimbabwe published by Zindi (1994) with the U.S. normative sample for Whites and found 77% of the between-groups race variance was attributable to a single source, namely *g*.

Another contribution of Rushton's was to confirm Spearman's hypothesis in South Africa by using *item analyses* of the Raven's Progressive Matrices. Since the Matrices are an excellent measure of *g*, it follows that each item's correlation with the test's total score provides a good estimate of that item's *g* loading. Using the method of correlated vectors, Rushton and Skuy (2000) examined 309 university students at the University of the Witwatersrand and found that the more an individual item from the Raven's measured *g* (estimated by its item–total correlation), the more it correlated with the differences in pass rates between Africans and Whites on the same items. In two studies of engineering students in South Africa, Rushton, Skuy, and Fridjohn (2002, 2003) found that the more the items from both the Standard and the Advanced Matrices loaded on *g*, the better they predicted the magnitude of the African–East Indian–White pass rate differences on the same items. The *g* loadings showed remarkable cross cultural generalizability. For example, those calculated for the East Indian sample predicted the African–White differences.

Rushton, Skuy, and Bons (2004) corroborated Spearman's Hypothesis on Black-White differences in South Africa using a multi-group confirmatory factor analysis (MGCFA). This was another important study because it established that finding "Jensen Effects" is not an artefact of using the method of correlated vectors. These results also refuted the related criticism that the Raven's Matrices do not have a high *g* loading among Africans.

A particularly bold study by Rushton calculated item heritabilities on 58 items from the Matrices from 152 pairs of twins from the Minnesota Study of Twins Reared Apart (MISTRA; Bouchard & McGue, 2003). These predicted differences in item pass rates among 11 diverse groups including the Roma Gypsies in Serbia, White Canadian and US twins, as well as East Asian, European, South Asian, Colored, and African university students in South Africa (overall mean $r = .40$, $p < .05$; Rushton, Bons, Vernon, & Čvorović, 2007). Most of the twins in this database had been separated early in life, reared in adoptive families, and reunited only in adulthood. The heritabilities were calculated in several ways, including using the correlation between the monozygotic twins reared apart. The results were corroborated using a variety of methods including creating parcels of items in order to provide a more reliable composite than any single item. As the heritability of the parcels increased, so did the magnitude of the group differences. Using the MISTRA sample, the items were sorted into six parcels of nine heritabilities and correlated with the parcels of pass rate differences between the Minnesota twins and all the other groups, including the Roma in Serbia, and the East Asian, European, South Asian, Colored, and Black university student differences in South Africa (mean $r = .74$; range = .55 to .93; $p < .0001$).

Strong inference is possible: (1) Genetic theory predicts a positive association between heritability and group differences; (2) culture theory predicts a positive association between environmentality and group differences; (3) nature + nurture models predict both genetic and environmental contributions to group differences; while (4) culture-only theories predict a zero relationship between heritability and group differences. These

results provide strong and reliable corroboration of the hypothesis that the cause of group differences is the same as the cause of individual differences, that is, about 50% genetic and 50% environmental (Rushton & Jensen, 2005, 2010).

3. Brain Size-IQ Relations

Back at Berkeley in 1981 when Rushton and I first discussed the issue of brain size and evolution, very little was known about how it might bear on racial differences in IQ. There was at the time, a controversial theory by the physical anthropologist Carlton S. Coon (1962) asserting that the major races differed in evolutionary age, supposedly having crossed what he termed the *Homo erectus-Homo sapiens* threshold at different points in evolutionary history. There were also said to be average differences in cranial capacity and brain weight, in relation to overall body size, among races, and even social classes. Although I raised these topics in an early review (Jensen, 1973), I mainly limited my considerations to the quantitative genetic and psychometric aspects of group differences. I hoped that in due course students of anthropology and related disciplines would bring their specialized knowledge, methods, and evidence to bear on the topics.

Rushton turned his eyes to looking for that evidence. In 1980 an autopsy study found a 100 gram brain weight difference between Blacks and Whites in 1,261 American adults aged 25 to 80 from Cleveland, Ohio, from which obviously damaged brains were excluded (Ho, Roessmamm, Strauthmann, & Monroe, 1980). The 811 European Americans averaged 1,323 grams and the 450 African Americans, 1,223 grams. This 100 gram difference remained significant after adjusting for age, stature, body weight, and body surface area. Reliable data on cranial capacity appeared shortly thereafter. It showed an East Asian advantage (Beals, Smith, & Dodd, 1984). This was a database of 20,000 skulls from 122 ethnic groups, which Beals et al. classified in terms of climate and geography. They found a 2.5 cm^3 increase in brain volume with each degree of latitude away from the equator. Skulls (sex-

combined) from East Asia averaged 1,415 cm^3; from Europe 1,362 cm^3; and from Africa, 1,268 cm^3.

Like me, Rushton considered brain size to be an especially important variable. People with larger brains have been demonstrated to average higher IQs than do those with smaller brains (Rushton & Jensen, 2005). Rushton reviewed this literature several times, most recently with his zoology colleague C. Davison Ankney (2009). Over two dozen studies with 700 participants had established a significant relation between brain size and IQ using magnetic resonance imaging (MRI) to measure brain volume. They found an overall correlation of .40 with IQ. This is a much higher correlation than the one of .20 typically obtained using external head size measures (such as head perimeter), although that correlation too is significant. The brain size-IQ relation has been found for people of all ages, both sexes, and diverse backgrounds, including African Americans. One reason put forth as to why larger brains are more intelligent is they contain more neurons and synapses making them more efficient at processing information.

Rushton also added to the primary literature on race differences in brain size by calculating cranial capacities from measures of external head size. He tracked down and analyzed several large anthropometric surveys which had been collated by NASA (of international air-force pilots), by the U.S. Army (of 6,325 military personnel), by the International Labour Office (of tens of thousands of workers in diverse countries), and by the U.S. National Collaborative Perinatal Project (of 35,000 children followed from birth to 7 years). Rushton and Ankney (2009) also summarized the world literature on race differences in brain size for the three major methods (autopsies, endocranial volume, and head size measures). These averaged, in cubic centimeters: East Asians = 1364; Whites = 1347; and Blacks = 1267. Since one cubic inch of brain matter contains millions of brain cells and hundreds of millions of synapses or neural connections, Rushton argued that race differences in average brain size help to explain their differences in average IQ.

4. The *r–K* matrix of life-history traits

In reviewing my own work on educability and group differences, Rushton seized upon a footnote I had written that "the three racial groups lie on a developmental continuum on which the Caucasian group is more or less intermediate [to East Asians and Blacks] (Jensen, 1973, p. 289)." He then accumulated and collated data for the three major races on numerous other categories of behavior beginning with Richard Lynn's (1977, 1982) finding of a higher mean IQ in East Asian populations (IQ = 106, 110) than in White populations for whom the tests had been developed and whose culture it reflected (IQ = 100). This was true even though the socioeconomic conditions of the East Asian nations were lower than for the majority White nations (at least at that time). These findings present serious difficulties for purely cultural (that is, completely non-genetic) explanations of the Black-White IQ gap. Since those early studies, Lynn and Vanhanen (2002, 2006) have greatly increased the evidence for higher IQs in East Asian countries (mean IQ = 106) and lower ones in sub-Saharan African countries (mean IQ = 70), with European countries intermediate (mean = 100).

Perhaps most innovative of all was Rushton's application of E. O. Wilson's (1975) *r-K* life history theory to human race differences. Rushton had amassed and organized data for about 60 biological and behavioral variables on the three major populations descended from East Asia, Europe, and Africa. In addition to brain size and IQ, the other variables included speed of maturation and longevity, personality and temperament, family stability and crime, and sexual behavior and fertility (see Table 1 in Nyborg, this issue). Rushton proposed that all these traits co-varied because they form part of a life history—a genetically organized suite of traits that evolved together to allocate developmental resources to the goals of survival, growth, and reproduction.

Rushton then applied this general biological theory to human group differences. He (1985, 1995) proposed that East

Asians were more *K*-selected than Europeans who were more *K* selected than Africans.

Rushton has described *r-K* theory in terms of the origin of parental care. Parental care, he argued, can be scaled from, at one end, a "fast" life history (the *r*-strategy) in which eggs and sperm are produced and simply discharged into the water (for example, in oysters) to a "slow" life history (the *K*-strategy) in which an egg is not only laid in a nest and food and care provided for the young (as birds). In mammals, the combined demands of gestation, delivery, production of milk, and protecting and physically caring for the young reach a peak (see Figure 1 in Nyborg, this issue). *K*-strategists, Rushton has argued, provide a lot of parental care. They have complex social systems and work together in getting food and providing shelter. *K*-strategists have more developed nervous systems and bigger brains but produce fewer eggs and sperm. The bigger an animal's brain, the longer it takes to reach sexual maturity and the fewer offspring it produces. Number of offspring, time between births, parental care, infant mortality, speed of maturity, life span, even social organization and altruism all fit together like pieces of a puzzle.

Rushton has shown that animal species that adopt the *K*-strategy, especially monkeys, apes, and humans, have large brains, are more intelligent, and live longer than *r*-strategists. *K*-strategists give their offspring a lot of care. They work together in getting food and shelter, help their kin, and have complex social systems. And that is why, according to Rushton, the *K*-strategists also need a more complex nervous system and a bigger brain.

One striking example Rushton has offered of race differences in *K*: Around the world the rate of dizygotic (i.e., two-egg) twinning is less than four per 1000 births among East Asians, eight among Whites, and 16 or greater among Blacks (a finding we discussed in 1981). And importantly, the tendency to produce dizygotic twins is heritable and mediated by sex hormones. Another example: Black babies sit, crawl, walk, and put on their clothes earlier than Whites or East Asians. And again, these differences have been found to be heritable. For example, children of mixed

Black–Japanese ancestry average a faster rate of skeletal development than do children of mixed Japanese–White ancestry, who average a faster rate of skeletal development than do children with two Japanese parents. For walking, East Asians, 13 months; Whites, 12 months; Blacks, 11 months. Blacks also have an earlier age of sexual maturity than do Whites, who in turn have an earlier age than do East Asians, whether measured by age of first menstruation, first sexual experience, or first pregnancy.

5. Race differences and human origins research

The current consensus view of human origins, the ''out-of-Africa'' theory, posits that *Homo sapiens* arose in Africa about 150,000 years ago and then expanded northward beyond Africa about 100,000 years ago, with a European–East Asian split about 41,000 years ago. Rushton (1995) has made the case that evolutionary selection pressures were different in the hot savanna, where Africans lived, than in the cold northern regions Europeans experienced, or the even colder Arctic regions where East Asians evolved. Thus, the further north the ancestral populations migrated out of Africa, the more they encountered the cognitively demanding problems of gathering and storing food, making clothes and shelters, and raising infants in winter environments.

Neural complexity and brain size have been increasing over the last 575 million years of evolutionary history in both vertebrates and invertebrates (Rushton & Ankney, 2009), which are related not only to increasing behavioral complexity (i.e., intelligence) but also, according to Rushton, to the *r-K* matrix of life-history traits, at least in mammals. As large brains evolved, they required more prolonged and complex life histories to sustain them. Large brains are also metabolically expensive, representing 2% of body mass but consuming 5% of basal metabolic rate in rats, cats, and dogs, 10% in rhesus monkeys and other primates, and 20% in humans.

One powerful way to test the validity of his extension of *r–K* theory is to examine, across a number of species, the variables hypothesized to covary. Rushton (2004) therefore examined the

relation between brain size and life-history traits in 234 mammalian species. There was enormous variation among the species sampled. The low end of the size continuum included the Madagascar hedgehog (body length = 185 mm; brain mass = 2 grams); at the high end was the African elephant (body length = 5,000 mm; brain mass = 4,480 grams). Sometimes, the data came from a single zoo specimen and at other times from a survey of a wild population. A principal components analysis across a single *r–K* life history factor had loadings such as: brain weight (.85), longevity (.91), gestation time (.86), birth weight (.62), litter size (.54), age at first mating (.73), duration of lactation (.67), body weight (.61), and body length (.63). This, Rushton asserted, shows that having a *K* strategy is the keystone that explains a set of observed behaviors that rest upon it. In support of his position, he demonstrated that the factor loadings remained high even after body weight and length were statistically controlled. Neither did the results vary if a principal axis factor analysis was performed instead of a principal components factor analysis. So it is the *K* strategy, not body size, he concluded, that provides the prime mover.

6. Rushton and the Galtonian Tradition

Rushton's research conclusions have generated no small measure of controversy. Going back to Galton's three personality qualities of genius, however, he has shown even greater measures of intellect, zeal, and especially persistence in tackling the most controversial issues facing not only behavioral science but society as well.

Anthropologist Henry Harpending (1995), a member of the U.S. National Academy of Science in reviewing Rushton's 1995 book, *Race, Evolution, and Behavior* in *Evolutionary Anthropology* wrote:

> Rushton's *Race, Evolution, and Behavior*...is an attempt to understand [race] differences in terms of life-history evolution....Perhaps there ultimately **will**

> be some serious contribution from the traditional smoke-and-mirrors social science treatment of IQ, but for now Rushton's framework is essentially the only game in town.

Linda Gottfredson (1996), in her review in *Politics and the Life Sciences,* wrote:

> ....*Race, Evolution, and Behavior* confronts us as few books have with the dilemmas wrought in a democratic society by individual and group differences in key human traits.

To these I can only add that Rushton's contributions to the sciences Galton initiated are truly worthy of their founder.

References

Beals, K. L., Smith, C. L., & Dodd, S. M. (1984). Brain size, cranial morphology, climate, and time machines. *Current Anthropology*, *25*, 301-330.

Bouchard, T. J. Jr., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology*, *54*, 4-45.

Coon, C. S. (1962). *The origin of races*. New York: Knopf.

Gottfredson, L. S. (1996). Review of *Race, evolution, and behavior: A life history perspective* by J. P. Rushton. *Politics and the Life Sciences, 15*, 141-143.

Harpending, H. (1995). Human biological diversity. [Review of the books *The Evolution of Racism*, *Human Biodiversity*, *The Bell Curve*, and *Race, Evolution, and Behavior*]. *Evolutionary Anthropology*, *4*, 99-103.

Ho, K. C., Roessmann, U., Straumfjord, J. V., & Monroe, G. (1980). Analysis of brain weight. I and II. *Archives of Pathology and Laboratory Medicine*, *104,* 635-645.

Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.

Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger.

Jensen, A. R. (2006). *Clocking the mind*. New York: Elsevier.

Jensen, A. R. (2011). Editorial. The theory of intelligence and its measurement. *Intelligence*, *39*, 171-177.

Lynn, R. (1977). The intelligence of the Japanese. *Bulletin of the British Psychological Society*, *30*, 69–72.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature*, *297*, 222–223.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations*. Westport, CT: Praeger.

Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality*. Augusta, GA: Washington Summit.

Miele, F. (2002). *Intelligence, race, and genetics: Conversations with Arthur R. Jensen.* Boulder, CO: Westview.

Nagoshi, C. T., Johnson, R. C., DeFries, J. C., Wilson, J. R., & Vandenberg, S. G. (1984). Group differences and first principal-component loadings in the Hawaii Family Study of Cognition: A test of the generality of "Spearman's hypothesis." *Personality and Individual Differences, 5,* 751-753.

Nichols, P. L. (1972). *The effects of heredity and environment on intelligence test performance in 4- and 7-year-old white and Negro sibling pairs*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Nyborg, H. (2003). (Ed.) *The scientific study of general intelligence*. New York: Elsevier.

Nyborg, H. (Part I in this book). In conversation with J. Philippe Rushton.

Rushton, J. P. (1985). Differential *K* Theory: The sociobiology of individual and group differences. *Personality and Individual Differences, 6*, 441-452.

Rushton, J. P. (1989). Japanese inbreeding depression scores: Predictors of cognitive differences between Blacks and Whites. *Intelligence*, *13*, 43-51.

Rushton, J. P. (1995). *Race, evolution and behavior: A life history perspective*. New Brunswick: NJ: Transaction.

Rushton, J. P. (1998). The "Jensen Effect" and the "Spearman-Jensen Hypothesis" of Black-White IQ differences. *Intelligence, 26,* 217-225.

Rushton, J. P. (2001). Black-White differences on the *g*-factor in South Africa: A "Jensen Effect" on the Wechsler Intelligence Scale for Children-Revised. *Personality and Individual Differences, 31*, 1227-1232.

Rushton, J. P. (2004). Placing intelligence into an evolutionary framework or how *g* fits into the *r-K* matrix of life history traits including longevity. *Intelligence*, *32*, 321-328.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235-294.

Rushton, J. P., & Ankney, C. D. (2009). Whole-brain size and general mental ability: A review. *International Journal of Neuroscience, 119*, 691-731.
Rushton, J. P., Bons, T. A., Vernon, P. A., & Čvorović, J. (2007). Genetic and environmental contributions to population group differences on the Raven's Progressive Matrices estimated from twins reared together and apart. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *274*, 1773-1777.
Rushton, J. P., & Jensen, A. R. (2003). African–White IQ differences from Zimbabwe on the Wechsler Intelligence Scale for Children-Revised are mainly on the *g* factor. *Personality and Individual Differences*, *34*, 177–183.
Rushton, J. P., & Jensen, A. R. (2010). Editorial. The rise and fall of the Flynn Effect as a reason to expect a narrowing of the Black–White IQ gap. *Intelligence*, *38*, 213-219.
Rushton, J. P. & Skuy, M. (2000). Performance on Raven's Matrices by African and White university students in South Africa. *Intelligence, 28,* 251-265.
Rushton, J. P., Skuy, M., & Bons, T. A. (2004). Construct validity of Raven's Advanced Progressive Matrices for African and non-African engineering students in South Africa. *International Journal of Selection and Assessment*, *12*, 220–229.
Rushton, J. P., Skuy, M., & Fridjohn, P. (2002). Jensen Effects among African, Indian, and White engineering students in South Africa on Raven's Standard Progressive Matrices. *Intelligence, 30,* 409-423.
Rushton, J. P., Skuy, M., & Fridjhon, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, Indian, and White engineering students in South Africa. *Intelligence*, *31*, 123–137.
Skuy, M., Schutte, E., Fridjhon, P., & O'Carroll, S. (2001). Suitability of published neuropsychological test norms for urban African secondary school students in South Africa. *Personality and Individual Differences, 30,* 1413-1425.
Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York, NY: Macmillan.

Wilson, E. O. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Harvard University.
Zindi, F. (1994). Differences in performance. *The Psychologist, 7*, 549-552.

Obituaries for Arthur Jensen

Detterman, D. K. (2012). Thank you, Arthur Jensen (August 24, 1923- October 22, 2012). *Intelligence, 41,* Issue 2, 141-143.
Flynn, J.R. (2012). Arthur Robert Jensen (1923-2012). *Intelligence, 41*, Issue 2, 144-145.
Lynn, R., (20012). Obituary, Arthur Robert Jensen, 1924–2012. *Intelligence, 41*, Issue 2, 146-148.
Sterling, Cooper (2012). Remembering a Scientific Pioneer – Arthur R. Jensen (1923-2012). *Occidental Observer*, November 5.
Woo, Elaine (2012). Arthur Jensen dies at 89; his views on race and IQ created a furore. *Los Angeles Times*; Obituaries. November 2.

A **collegial tribute to Arthur Jensen**

Nyborg, H. (2003; Editor). *The scientific study of general intelligence: Tribute to Arthur R. Jensen.* Oxford: Pergamon/Elsevier Science. (642 p.)

Chapter 2

Resolute Ignorance on Race and Rushton

Linda S. Gottfredson
School of Education, University of Delaware
Newark, DE 19716 USA.

Abstract

I review J. P. Rushton's research on the evolutionary divergence of the three major human lineages. His life-history theory predicts, and his multiple analyses document, a consistent three-way patterning of mean differences among blacks, whites, and East Asians on coevolved sets of morphological, physiological, developmental, psychological, and behavioral traits. I then analyze a typical example of how critics evaluate his work, including the rate at which they cast his scientific hypotheses, methods and conclusions in politically charged language. The set of articles in question, although authored by well-known academics and appearing in a major, peer-reviewed journal, illustrate how mob science works to "discredit" valid research and enforce collective ignorance about entire bodies of evidence. Rushton is a scholar and gentleman but it appears that his critics often act like neither.

Introduction

J. Philippe (Phil) Rushton has contributed important works to evolutionary psychology, intelligence, and personality psychology. I focus here on his work receiving the most attention. That is his r-K theory of how the major races, or geographic lineages, evolved in somewhat different directions as humans spread out of Africa about 50-100mya. He has documented a large suite of morphological, physiological, developmental, psychological, and behavioral differences among these groups, whose most recent common ancestors are from Africa, Europe and East Asia. These various traits cohere evolutionarily and are consistent with his r-K life history explanation of the wide array of mean group differences that persist over generations.

I first assess Rushton's research contributions. Then I analyze a typical example of the scornful commentary on the man and his work—a target article and eight comments published together in a major, peer-reviewed journal. I look especially at the nature of evidence and argument used by Rushton and the authors asked to evaluate his work.

1. Rushton's approach to human biodiversity

Rushton is proudly of the London School of thought in psychology. It rejects the separation of mind from body and of culture from its genetic substrate, preferring instead to probe their connections. It was an outpost of biological realism during the long reign of behaviorism, whose founder famously assured us we could form children into anything we wished, not unlike the Soviets' New Man.

Phil Rushton began his career by addressing one of evolutionary psychology's biggest challenges at the time. Altruism seems to require reproductive self-sacrifice, so how could it possibly have evolved? This work garnered him praise and a Guggenheim Fellowship. Humans are not promiscuous altruists, of course, but favor persons genetically similar to themselves. This led

him to ponder the dissimilarities that have intrigued writers and travelers throughout human history: Why do the different tribes of man look and live so differently?

During the 1980s Rushton began systematically testing a theory-based life-history explanation. As I describe later, the life-history perspective allowed him to predict a particular pattern of evolved differences among genetic lineages. These predictions are not obvious because they link seemingly unconnected attributes across different realms of human existence, from sexual behavior to social organization. The great sweep yet high specificity of his theory with regard to racial differences would seem to make it easy to disprove if false. To test it, Rushton collected a broad spectrum of primary and secondary data, comprising three categories.

1.1 Evolutionary life history

Life history is a population-level concept. It refers to the coordinated suite of traits and behaviors that characterizes a particular species or subspecies (its shared "life"), which it evolved in response to the recurring adaptive challenges its members faced (its evolutionary "history"). The life-history concept highlights an important empirical phenomenon. A species' distinctive traits—such as humans' large brain, slow maturation, and pair bonding—do not evolve independently, one by one, but as a constellation of co-evolving traits. In fact, the nexus of traits typifying a species had to have evolved in a coordinated manner for the organism to maintain or enhance its evolutionary fitness. For instance, a bigger brain relative to body size consumes more of an organism's fixed energy budget, so humans (those who survived) evolved a smaller gut to afford their metabolically expensive brains, which in turn required a less toxic and more easily digested diet, which the innovation of cooking provides. The physical and logistic constraints of bearing and nurturing big-brained babies likewise generated correlated selection pressures on social behaviors, including male and female mating strategies, pair bonding, parenting, and provisioning.

Rushton relies on the r-K version of life history theory that evolutionary biologists have used to distinguish species throughout the animal kingdom. The r-K version arrays organisms along a continuum of reproductive strategies, from highly r-selected species (many offspring, little parenting) to highly K-selected species (few offspring, much parenting). The assumption is that, given its evolutionary importance, reproductive strategy is likely the axis around which other species-typical traits become organized. The two strategies, r and K, are alternative means to the same end: reproductive fitness, which is to produce more genetic descendants than the Joneses.

Biologists have used r-K theory to compare members of the same species, but Rushton was the first to apply it so systematically to humans. His aim has been to determine whether, and how well, it can explain the many systematic differences observed today among major branches of the human family. He focuses on three. While their labels have changed over time, their origins have not: lineages tracing their ancestry to sub-Saharan Africa (Negroids or blacks), East Asia (Orientals or East Asians), and the lands in-between (Caucasians or whites).

Rushton gathered already-published population data on several dozen traits from institutional sources such as the US military (e.g., head size) and Interpol (international crime rates). His most recent update (Rushton & Jensen, 2005, Table 3; reproduced in Nyborg, **Part I**, *In conversation with J. Philippe Rushton* in this book) compares the three geographic lineages on 26 measures: intelligence (3 indicators), brain size (2), maturation rate (7), personality (5), reproduction (6), and social organization (3). The set forms a consistent pattern illustrating key insights from life history theory, in turn supporting his application of it to humans.

First, mean population differences are numerous, consistent, and generalized across the body and behavior. Whites are intermediate to blacks and East Asians on virtually all (24) of the 26 measures, whether they be physical or behavioral, speed or size. For instance, physical maturation (skeletal, motor, dental, lifespan) and sexual maturation (age of first intercourse, first

pregnancy) are accelerated in blacks relative to whites, and in whites relative to East Asians. Compared with whites, American blacks have larger bodies (reported elsewhere), larger secondary sex characteristics, higher hormone levels, and higher rates of sexual intercourse and two-egg DZ twinning (not MZ twinning)—with the reverse being true for reproductive behavior in East Asians compared with whites. A three-way pattern of mean differences is also seen in personality, with blacks being the most (and East Asians the least) aggressive, impulsive, dominant and sociable, whereas East Asians are the most cautious (and blacks the least) on average. It is also seen in social organization, with East Asians having higher rates of marital stability, law abidingness, and mental health than whites, and whites having higher rates than blacks.

The apparent evolutionary divergence between the three lineages is not just from "the neck down." East Asians have the largest skulls, largest brains, and most cortical neurons, whereas blacks have the smallest skulls and brains and the fewest cortical neurons, on average. These differences in the physical brain are mirrored by comparable mean group differences in the speed and quality of the brain's information processing. East Asians have the fastest decision reaction times (measured in milliseconds) on elementary cognitive tasks, which are so simple that virtually everyone can perform them correctly, and they also have the highest average levels of general intelligence (*g*) on validated, unbiased tests of intelligence.

The pattern of cranial vs. sub-cranial group differences illustrates a second insight from life-history theory: evolution imposes tradeoffs. For example, the populations with larger brains have (had to evolve) smaller bodies; and those with slower maturing offspring have (had to evolve) more socially and sexually constrained adults.

Note that none of the aforementioned biological traits is a social marker of race, either singly or collectively, as would be skin color and hair type. Moreover, as Rushton and others have shown, within-race variation in all these characteristics is moderately to highly heritable, so we cannot dismiss the possibility that mean

group differences in body and behavior are also somewhat genetically rooted. In fact, these mean differences in "non-racial" traits appear to shift in tandem, as a coordinated set, from one human lineage to another. This implicates a consistent deep influence linked more tightly to distant genetic heritage than to current circumstance. The great variation within racial groups is entirely consistent with genetic divergence, because within-population variation is the grist for evolution. Also consistent is the systematic overlap among groups, because mean differences in genotype will emerge from the same ancestral genotype, slowly but surely, when the groups' adaptive demands diverge and consistently pull selection in somewhat different directions.

Rushton's results reflect a third insight of life history theory: individuals do not evolve, populations do. A population's social organization—its culture—necessarily co-evolves with the distribution of its members' attributes. Humans, for example, are not just an exceptionally brainy primate, but also an especially social one. We have a special penchant for pair bonding (even if serial), living and working in groups, forming networks and coalitions, trading, teaching, gossiping, reading others' minds, and befriending non-kin. Our big-brained, slow-maturing, vulnerable and care-intensive children would not survive without such enduring social networks and bonds of long-term reciprocity.

Rushton's r-K version of life-history theory proposes that reproductive strategies drive (cause) differences in social organization. Although all humans are K-oriented (relatively few offspring with much parenting), his r-K theory posits that the somewhat less K-oriented human populations will tend more strongly toward social relations guided by self-interest, relative physical strength, and a tolerance for interpersonal conflict, with the result that, from a K perspective, they produce cultures with less stable families, less organized institutions, more transgressions of person and property, and therefore higher rates of mental and physical illness. In contrast, more K-oriented lineages produce societies whose members tend to exercise more self-control, social control, and mutual coordination in the pursuit of longer-term

shared goals, with the result that such groups will create somewhat better organized, more productive, personally secure and—from an r perspective—more rigid societies that tightly constrain what their members may do. Non-evolutionary explanations have been offered for these differences in social organization but none, to my knowledge, can account for—let alone has predicted—the nexus of physical, mental, and behavioral mean differences that Rushton found, and had predicted.

1.2. Forensic anthropology

Rushton has also tested his life-history theory by following hominoids back down their evolutionary tree. Rushton and Rushton (2004) examined progressive changes in 76 musculo-skeletal traits across seven hominoid populations, listed here by evolutionary age: *Pan troglodytes* (chimpanzee), *Australopithecines, Homo habilis, Homo erectus,* Africans, Europeans, and East Asians. The aim was to test the hypothesis that a cascade of skeletal changes accompanied the evolution of brain size in hominoids, ranging from humans' most distant relative (the chimpanzee, 5 mya) to the youngest human lineage (East Asians). Standard texts on evolutionary anatomy provided data on 76 skeletal traits for the one ape and the three fossil species, and standard forensic anthropology textbooks provided data on 42 traits for the three human populations.

The skeletal data for the seven hominoid groups did, in fact, differentiate them in the same order as did their evolutionary distance from the youngest human lineage (East Asians). Of the 42 traits available for the three racial groups, 38 were measured in absolute terms—14 cranial traits (including cranial capacity), 8 teeth and mandibular traits, 3 neck, 3 pelvic, and 12 upper and lower limb traits. All but one (nasal bone prominence) fit the predicted pattern.

The 6 other traits, measured as body proportions, uniformly did not (e.g., leg length as a % of height, upper arms as % of body weight). Ratios and percentages of a trait are hard to interpret because they have different measurement properties than do

absolute measures of a trait, such as length, area, volume and weight. Nor is it clear that r-K theory makes any predictions for body proportions. If we set the body proportions data aside for now, it appears that the increase in cranial capacity from 380 to 1364 cm^3 across the seven hominoid groups was accompanied by systematic and pervasive changes across the musculo-skeletal system, including cranial traits (e.g., size and shape of the skull, jaw, teeth, eye sockets, brows, muscle attachments) as well as post-cranial traits (e.g., dimensions, shape and orientation of particular bones, joints, and pelvis). The most general change is that the musculo-skeletal system became less robust as brain size increased. Perhaps there is a non-evolutionary explanation for the progressive and pervasive skeletal differences among blacks, whites, and East Asians so well known to forensic investigators, but I am not aware of any plausible ones having been offered.

1.3. Patterns of phenotypic and genetic variation in intelligence

If the major human races diverged physically and behaviorally during evolution, their living descendants should differ genetically, on average. Self-identified races are different branches of the human family, as confirmed in the last decade by studies of the Y chromosome, mitochondrial DNA, neutral markers on the 22 pairs of autosomal chromosomes, number of short tandem repeats on various ones, and more. But that is not the issue here. It is whether a highly particular, much documented, mean phenotypic difference among the major human races has a genetic component.

Of all human traits, variation in general intelligence (*g*) is the functionally most important in modern life. The first question that behavior genetics tackled was "how heritable are within-group differences in intelligence?"—the answer: "very." The next obvious question is "how heritable are the between-group differences in phenotypic intelligence?" It could easily be answered using today's analytical methods, but no scientific discipline will touch it. Most would have to be shut down completely, however, to avoid generating relevant evidence. Traits that are so useful in daily life and so heritable across generations are going to manifest

themselves in many predictable ways. Rushton has marshaled relevant such evidence on intelligence to test competing hypotheses about the long-standing mean IQ difference between American blacks and whites: 0% genetic vs. 50%-80% genetic.

The latter is Jensen's (1998) "default hypothesis," which is that within-race variation and between-race variation arises from the sources, whether genetic or environmental. Accordingly, there is no Factor X operating on all members of one race but on no member of another. (Anything that affects some but not all members of a group would show up as a within-group influence.) If within-race IQ variation is 50%-80% heritable, as it is in the West, the default hypothesis predicts that between-group differences will be too. This is a readily testable hypothesis, but virtually no one with the necessary data has been willing to test it or lend the data to others who are.

Rushton and Jensen (2005) tackled the question by having their "hereditarian" hypothesis (50%-80% genetic) compete head-to-head with the "culture-only" hypothesis (0% genetic) in 10 categories of evidence: the worldwide distribution of test scores, the *g* factor of mental ability, heritability, brain size and its relation to mental ability, transracial adoption, racial admixture, regression to the mean effects, related life-history traits, human origins research, and hypothesized environmental influences on intelligence. This set captures much of the dense nomological network of empirical evidence on psychometric *g*, including its genetic structure, biological and social correlates, behavioral manifestations, joint heritability with brain structure, function, and life outcomes, developmental course, and manipulability by various experimental means, including adoption.

The evidence in at least 7 of the 10 categories is much replicated, often over time, age, sex, race, class, and geography, and therefore provides a firm basis for comparison. The 50%-80% genetic hypothesis fits this array of evidence far better than does the culture-only hypothesis (Gottfredson, 2005). Where the two make opposite predictions, the hereditarian predictions are confirmed and the culture-only predictions are contradicted. For

instance, achievement differences expand just when the culture-only explanation says they will contract—when resources and opportunity are equalized (Ceci & Papierno, 2005). And, as noted earlier, group differences follow geographic ancestry far more closely than the groups' current location or circumstances. They behave just like evolved, genetically-influenced traits are expected to behave—with great consistency regardless of species-typical variations in non-genetic milieu. The greater explanatory success of the 50%-80% genetic hypothesis does not prove it true, but that is better approximates the truth than does its never-plausible 0%-genetic competitor.

2. The critics' approach to Rushton and race

In 1989 Rushton summarized his evidence for a black-white-East Asian gradient in life-history traits at the AAAS meeting in San Francisco. There followed a convulsion of excoriation by the scientific establishment, his home institution (University of Western Ontario), and the Canadian government (Rushton, 1998). All launched investigations into his work. He was shunned and isolated by fellow academics, as if having a professionally deadly contagious disease. When he presented his corpus of published evidence in book form—*Race Evolution, and Behavior: A Life History Perspective* (1995)—it was greeted as "inflammatory," "indecent," "pseudoscientific," "racist trash," and in scatological terms too (Gottfredson, 1996).

2.1. High talk and low blows

In a collective exercise in confirmation bias, Rushton's critics spied damning evidence of scientific and moral perfidy wherever they looked. They saw nothing exculpatory, or even ordinary, in his conduct. So, where evolutionary psychology had always concentrated on sexual selection and therefore on mating strategies, sexual behavior, and physical attractiveness to the opposite sex, critics insinuated that Rushton's interest was prurient. Although behavioral geneticists were busy estimating the heritability of

intelligence differences within races, Rushton's interest in possible between-race genetic differences betrayed, to them, an archaic racial elitism. And while none denied that the human species' remarkably large brain is responsible for its remarkably high intelligence, they said Rushton was resurrecting long-discredited 19th century thinking when he asserted (correctly) that brain size and intelligence are correlated in modern humans.

Rushton answered all published critiques while carefully adhering to the scientific coda his critics often flagrantly violated—logic, weight of evidence, and no aspersions on character. Rushton's dispassionate scientific manner on socially sensitive questions only confirmed for them that he was a heartless ideologue misusing science for pernicious ends. I illustrate this phenomenon of high talk and low blows, as Robert Gordon calls it, with a concrete example. It is a target article (Lieberman, 2001) and associated comments published in *Current Anthropology*. The *Anthropological Review's* 2007 obituary for Lieberman lauded him for having "often challenged racists and racialist views." It pointed specifically to his 2001 article because it "dissected and ridiculed [Rushton's] views of supposed racial differences in intelligence."

2.2. Argument from authority, political opinion, and impossible standard of proof

Lieberman opens with a question that itself damns Rushton. A century of anthropological work has invalidated Rushton's claims, so how can he claim to find in it a "racial hierarchy" for intelligence and brain size? The question is thus not whether Rushton is wrong, but why and how he persists in being so wrong. The article's first section ("Changing Hierarchical Worldviews") justifies the premise, and the second ("Abusing Anthropological Research") explains the "paradox" of how Rushton and other "scientific racists" could claim to be doing science when they draw evidence from the very fields that disavow racism and the concept of race (p. 74).

To justify his premise, Lieberman describes 19th century research on cranial size and its social context which, he says, was

the need by Caucasians to justify their domination and exploitation of other races. He discredits that research and Rushton's own primarily by appealing to authority in 20th century anthropology: Franz Boas's theorizing (no link between culture and genes), official statements on race from the UN and the American Anthropological Association (no biological races, no meaningful innate racial differences), Gould's critiques of research on intelligence, brain size, and heritability (none is valid), and anthropology's "disavowal" of "hierarchical" and "racist" thinking. Lieberman also draws from stock concerns, long-since resolved, about possible methodological flaws in twin research, brain research, behavior genetics, and mental testing. He says nothing about the explosion of research in the 1990s using the Y chromosome and mtDNA to trace the evolution of human lineages as they migrated across the globe. He says nothing about Rushton's many other 3-way results on "reproductive behavior, sex hormones, twinning rate, speed of physical maturation, personality, family stability, law-abidingness, and social organization" (p. 74), except to summarily dismiss them as a faulty "blizzard of data" (p. 78).

Zeroing in on Rushton's IQ, race, and brain size analyses, Lieberman details his own list of six major "errors." Briefly, Rushton "uses 'race' despite decades of findings that invalidate it," his conclusions about racial differences in cranial capacity are "contradicted by evolutionary anthropology," he did not account for environmental factors that surely influence cranial capacity and intelligence, his measurements tell us nothing because they are confounded or the differences they reveal are trivial in size, he cannot claim to "explain" a vast array of human behaviors because some of his measures and concepts may be faulty, and his "principle of aggregation" (e.g., grouping diverse populations into "races," averaging results from different studies) is invalid. "Aggregation" is Lieberman's single most frequent complaint of the six (38 times in 14 pages), even though averaging non-comparable samples would work against Rushton's finding consistent patterns.

The first two "errors" are, again, appeals to authority. The next two disallow drawing conclusions until an infinite regress of

alternatives has been considered, and the last two demand uniformly perfect data and measurement before concluding anything from a body of evidence. All insulate his factual premise (no evolved differences) from disconfirmation by creating five thou-shall-nots that selectively handicap researchers who might disagree. Scientists shall not refer to race except as a social construct; not group people or results by race unless to illustrate environmental effects; not hypothesize evolutionary differences that contradict reigning "antiracist" opinion in anthropology; not infer any genetic differences before ruling out all non-genetic influences; and not draw non-reigning conclusions about patterns of racial differences in a large body of data if any particular datum might be faulty.

Lieberman moves Rushton's work into the political realm by always labeling his hypotheses and results with political terms. It is "hierarchical" (because it measures ordinal differences on a trait), "racist" (it finds mean racial differences in traits), and a "justification" for "inequities" (it predicts social inequality when those traits matter). Table 1 illustrates this practice by categorizing his terms of derogation. (Data for detailed categories are available from the author.) He uses an average of 21 per page, split evenly between connoting extreme right-wing politics and unacceptable science. Forty percent (8.4 per page) allege error in politically neutral terms ("lacks evidence"), but almost as many render his hypotheses and results politically noxious by relabeling them hierarchical and racist (4.9 and 2.8 times per page).

The lone dissenting commentary, by Henry Harpending, objects to Lieberman repeatedly attributing notions of "inferiority" and "superiority" to Rushton (56 times, not shown), but Lieberman replies that Rushton's work implies them. He further connotes immoral politics by describing Rushton's research as "notorious," "destructive," and "socially harmful" and situating it within a history of political evils (slavery, genocide, torture, exploitation) and social harm (misery, poverty, inequality).

Table 1. Terms used to impugn Rushton without rebutting his evidence.

Category	Lieberman (2001) 12 pages of article, 2 of reply	Associated 6 hostile commentaries (7 pages total)	Examples (Italicized items appeared only in Lieberman, boldfaced only in commentaries)
	Instances per page		
Evil Ideology			
Hierarchical thinking	4.9	4.4	Essentialism, **Eurocentric, Caucasoid-centrism,** hierarchical or typological thinking (many variations)
Racist thinking	2.8	7.1	Racism (**latent, academic, scientific**), **racial ideology,** inferiority**, racial profiling**, xenophobia
Evil people and events	0.7	0.9	**Apartheid,** colonialism, genocide, **Holocaust**, *Nazi, Jim Crow segregation, slavery*, Southerners
Destructive results	0.3	1.1	*Destruction of social generosity*, **discriminatory, downtrodden**, inequity (& variations), misery
Subtotal	10.0	16.1	
Unacceptable science			
Incompetent	8.4	6.0	Aggregated, confounded, deterministic, erroneous, **illogical**, lacks (evidence, validity, *conceptual and empirical merit*), *seriously flawed*, **statistical artifact**, *substandard*, **subjective, unsophisticated**

Pseudoscientific	2.0	5.7	**Absurd, bad biology**, biased, **contrived, deceptive, diseased, fabricated (races)**, fascinated by sexuality of savages, **fetishizes the brain, inexcusable anthropology**, *mischievous speculation*, **reactionary, same old lies, self-serving, sorry mess**
Dangerous	0.4	0.9	**Alarming, dangerous**, *destructive, notorious*, **odious**
Subtotal	10.8	12.6	
Total per page	20.8	28.7	
(N instances)	(291)	(201)	

2.3. Mob science

Rushton was invited to submit a comment, as was one anthropologist "well-known [for his] support of racial differences" (p. 90). The other six individuals were already on record as hostile to such ideas. They are less restrained than Lieberman, their negative descriptors being more numerous (29 vs. 21 per page) and more extreme ("odious," "quackery," "same old lies"). They less often use politically neutral terms to allege scientific incompetence, but are over twice as likely as Lieberman (per page) to associate Rushton with racist thinking, evil politics, pseudoscience, social harm, and imminent danger. Their commentaries vary in emphasis and personal abuse, as such commentaries usually do: C. Loring Brace ("inexcusable anthropology"), Fatimah Jackson ("diseased," "twisted," "same old misrepresentations"), Jonathan Marks ("modern creationism," "quackery"), John Relethford ("resurgence of racial classification"), Audrey Smedley ("so-called science"), Verena Stolke ("continuity of racist thought," "persisting exclusions"), and Fredric Weizmann ("strong claims," "relationships of minimal importance").

Taken as a whole, the symposium illustrates what happens when high talk and low blows is practiced collectively: unrestrained mob action to destroy a purportedly vile member of the group, invite a single defender to speak from the sidelines, and allow the target to say a few words which the crowd will ignore or ridicule.

2.4. Resolute ignorance

The first rule in science is to consider the totality of evidence; the second is to make alternative hypotheses compete in explaining it. Rushton has done both, but Lieberman and commentators do neither. Table 2 helps illustrate how a hostile crowd can circumvent these rules yet still appear scientific in order to maintain "resolute ignorance" about some stubborn, unwelcome fact, as the late William Beer dubbed it—in this case mean racial differences in general intelligence.

The table lists the seven common rebuttals, ranging from "intelligence doesn't exist" to "racial differences are unthinkable." Most critics accept some of the foundational findings (second column) but seldom the same ones, meaning their "yes-buts" often clash. One may "discredit" the notion of racial gaps in intelligence by first accepting some of the evidence ("Yes, intelligence exists") but then reject the next link in the chain of evidence ("but it can't be measured fairly"). Another may concede that "Yes, it can measured," but reject a different link in the evidentiary chain ("but it isn't important in real life"), yet both stand arm-in-arm to denounce the evidence. All that matters in mob science is that critics howl together at the target.

Lieberman and fellow critics likewise jab haphazardly at different nodes in Rushton's network of evidence. All dismiss his hypothesis of evolved racial differences in intelligence on the grounds that races don't exist. But the evidence does not melt away for being relabeled, ignored, or called nasty names. Therefore, in "yes-but" fashion, some of the seven add that intelligence doesn't exist either; others that it's not important, or it's not as heritable as it seems; yet others that the race-IQ gap is trivial, or will be washed away by the Flynn Effect, or Gould discredited the whole business

of measuring intelligence and brain size. They reunite again in suggesting that no credible scientist could possibly agree with Rushton. Yet it is Gould's work on cranial capacity, not Rushton's, that we now learn was fudged and falsified (Lewis et al., 2011)—just as Rushton said it was.

Table 2. "Yes-buts" commonly used to ignore evidence on mean racial differences in intelligence (g)

Gambits for ignoring links in chain of evidence *"Racial gaps in intelligence are .."*	**"Yes"** A bedrock finding is conceded.	**"But"** A false claim is pressed to justify ignoring the bedrock finding	**"Because"** Fallacious "evidence" used to support the false claim. *Examples for*: **D**isproved hypotheses **I**rrelevant truths **E**motional appeals
1. Nonexistent	Yes, there is a black-white gap in IQ scores	But, there is no such thing as "intelligence" (or *g*)	Why? Disagreement over its definition (I); Only a cultural (or statistical) artefact (D); Undemocratic (E)
2. Mismeasured	Yes, IQ tests do measure (developed) intelligence	But, tests culturally biased against blacks	Why? Tests created by whites (I); Require cultural knowledge (I); Ruse to justify discrimination (E)
3. Unimportant	Yes, black-white IQ gap reflects gap average intelligence	But, it's trivial or lacks practical importance	Why? Important only when treated as such (D); Narrow academic ability (D); Not a measure of human worth (I, E)
4. Malleable	Yes, intelligence has practical importance	But, intelligence is malleable	Why? Brains, skills and knowledge change with age and experience (I); Must not "give up" on less able persons (E)

5. Environmental	Yes, differences in intelligence are stubborn (resistant to intervention)	But, intelligence is not genetically determined	Why? Abilities not "fixed" (I); Genes and environments interact (I); Differences explained by social advantage (D)
6. Disproved	Yes, IQ differences within a race are quite heritable	But, the average differences between races have been proved not to be genetic	Why? There is no gene for race (I); "Race" is a social not biological phenomenon (D); Racists are hereditarian (E)
7. Unthinkable	Yes, a 50%-80% genetic black-white IQ gap fits the evidence	But, we must not let anyone suggest the gap is partly genetic	Why? Dangerous if hypothesis is false (E); Even more dangerous if true (E); Hitler thought the gaps genetic (E)

References

Ceci, S.J., & Papierno, P.B. (2005). The rhetoric and reality of gap closing—when the "have-nots" gain but the "haves" gain even more. *American Psychologist, 60*, 149-160.

Gottfredson, L.S. (1996). Review of *Race, evolution, and behavior: A life history perspective*, by J.P. Rushton. *Politics and the Life Sciences, 15*, 141-143.

Gottfredson, L.S. (2005). What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law, 11*, 311-319.

Jensen, A.R. (1998). *The g factor*. Westport, CT: Praeger.

Lewis, J.E., DeGusta, D., Meyer, M.R., Monge, J.M, Mann, A.E., & Holloway, R. (2011). The mismeasure of science: Stephen Jay Gould versus Samuel George Morton on skulls and bias. *PLOS Biology, 9*(6):e1001071.

Lieberman, L. (2001). How "Caucasoids" got such big crania and why they shrank: From Morton to Rushton. *Current Anthropology, 42*(1), 69-80.

Rushton, J.P. (1998). The new enemies of evolutionary science. *Liberty, 11*(4), 31-35.

Rushton, J.P. (1995). *Race, evolution, and behavior: A life history perspective*. New Brunswick, NJ: Transaction Publishers.

Rushton, J.P., & Jensen, A.R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*(2), 235-294.

Rushton, J.P., & Rushton, E.W. (2004). Progressive changes in brain size and musculo-skeletal traits in seven hominoid populations. *Human Evolution, 19*(3), 173-196.

Chapter 3

The Flynn Effect, Group Differences, and *g* loadings

Jan te Nijenhuis
University of Amsterdam
Work and Organizational Psychology, University of Amsterdam, Amsterdam, the Netherlands.

Abstract

Flynn effect gains are predominantly driven by environmental factors. Might these factors also be responsible for group differences in intelligence? Group differences in intelligence have been clearly shown to strongly correlate with *g* loadings. The empirical studies on whether the pattern of Flynn effect gains is the same as the pattern of group differences yield conflicting findings. We present new evidence on the topic using a number of datasets from the US and the Netherlands. Score gains and *g* loadings showed a small negative average correlation. The general picture is now that there is a small, negative correlation between *g* loadings and Flynn effect gains. It appears that the Flynn effect and group differences have different causes.

1. Introduction

Secular gains in IQ test scores are among the most intriguing and controversial findings in psychology. James Flynn (1984) was the first to show that average scores on intelligence tests are rising substantially and consistently, all over the world. Between 1930 and 1990 the gain on standard broad-spectrum IQ tests averaged three IQ points per decade. For verbal tests, or more precisely, tests with a content that most reflects the traditional classroom subject matter, the gain is 2 IQ points per decade, and for non-verbal (Fluid and Visual) tests 4 IQ points per decade. Gains on specific measures, such as the Raven's Progressive Matrices when used for the assessment of military recruits average about 7 IQ points per decade.

Group differences in mean IQ are more the rule than the exception (Lynn & Vanhanen, 2002). Large-scale research shows that, on average, the adult American Black population scores below the White population by about 1.2 standard deviations, equivalent to about eighteen IQ points (Jensen, 1998). There is some discussion about the gap diminishing (see Dickens & Flynn, 2006ab; see Rushton & Jensen, 2005). Dutch-first-generation non-Western immigrants differences are about the same size as the US Black/White differences, but become substantially smaller for the second generation of non-Western immigrants (te Nijenhuis, de Jong, Evers, & van der Flier, 2004). Jensen (1998, pp. 380-383) has shown that *g* loadings correlate about .60 with Black/White IQ test score differences. Te Nijenhuis and van der Flier (2003) showed that Dutch-non-Western immigrant differences are also strongly predicted by *g* loadings.

What, then, are the causes of these differences? Some argue there is a strong genetic component to group differences (Rushton & Jensen, 2005), whereas others argue group differences are wholly caused by the environment (Nisbett, 2009). The secular gains are massive and the time period too short for large genetic changes in the population, so therefore the changes must be largely environmental – although reduced inbreeding has been suggested

to play a role (Mingroni, 2007; see also Woodley, 2011). It is an empirical question whether the strong environmental forces causing the scores over generations to rise are the same as the forces causing the group differences.

1.1. Rushton's Contribution to the Area

Rushton (1989) joined the fray on this topic showing that inbreeding depression scores from Japan predicted the magnitude of the Black/White differences on the same subtests in the US. Inbreeding depression is an established genetic phenomenon that occurs when people who are genetically related have children together, thereby producing in their offspring, on average, a lower score on IQ than would otherwise have been the case.

Subsequently, in an exchange with Flynn, Rushton (1999) showed that secular gains from the US, Germany, Austria, and Scotland had modest to small negative correlations with *g* loadings. This is an important result given that *g* loadings correlate substantially with group differences, as shown by Rushton in a series of articles. In South Africa *g* loadings of items of the Raven Matrices predicted mean differences on the items between White, South Asian, and Black students (Rushton, Skuy, & Bons, 2004; Rushton, Skuy, & Fridjohn, 2002, 2003). In Zimbabwe *g* was a strong predictor of the score differences between African and White 12- to 14-year-olds on the WISC-R (Rushton & Jensen, 2003). In Serbia item *g* loadings from the Raven Matrices predicted mean differences between Roma and Whites (Rushton, Čvorovič, & Bons, 2007).

These findings increase in importance when one takes into consideration that most studies show that *g* loadings of tests correlate highly with their heritabilities (Jensen, 1987; Pedersen, Plomin, Nesselroade, & McClearn, 1992; Rijsdijk, Vernon, & Boomsma, 2002; Spitz, 1988). Moreover, Rushton, Bons, Vernon, and Čvorovič (2007) computed heritabilities for items from the Raven's Standard Progressive Matrices and showed that they could predict various group differences.

Rushton (1999; Rushton & Jensen, 2010) argues that if the gains over generations had been similar to the B/W differences they would have been in line with environmental explanations of group differences; when the pattern in secular score gains is not similar to the pattern in B/W differences, the latter is in line with a genetic component in group differences. However, it should not be forgotten that these findings are at best indirect evidence.

Rushton's (1999) finding has been challenged by Flynn (2000) and Nisbett (2009) who claim that there actually is a substantial positive correlation between secular score gains and *g* loadings. If this were a fact, it would indeed jeopardize Rushton's position; it would mean that *g* loadings correlate highly with both environmental and genetic effects, making them useless. Since Rushton's study suggesting secular trends are not related to *g*, various other studies have been carried out (Colom, Juan-Espinosa, & García, 2001; Flynn 1999, 2000; Must, Must, & Raudik, 2003; Wicherts et al., 2004) yielding conflicting findings.

Flynn (2007, 2010) states that even if there is a small negative correlation between secular score gains and *g* loadings it sheds no light on the race and IQ debate. Flynn (2010) accepts the empirical findings that Black/White score differences on subtests of IQ batteries rise as their *g* loadings, cognitive complexity, heritability, and inbreeding sensitivity rise. However, he argues that the fact that the performance gap is larger on more complex tasks than on easier tasks does not necessarily tell us something about genes versus environment. For instance, he hypothesizes that when one group has better genes for height and keen reflexes, but finds itself in a less rich basketball environment – less incentive, low-quality coaches, less play – the environmental disadvantage will expand the between-group performance gap as complexity rises, just as much as a genetic deficit would. The skill gap between challenged and unchallenged players is hypothesized to be more pronounced the more difficult the task. So, someone exposed to an inferior environment will hit a "complexity ceiling" and this ceiling does not differentiate whether the phenotypic gap is due to genes or environment. Elsewhere Flynn (2007; see also Dickens & Flynn,

2001) has argued that Blacks tend to be systematically underexposed to cognitive complexity throughout their life-courses. Flynn argues that the correlations reported by Rushton do not decide the causal question.

1.2. Research Question

We used a number of datasets to see whether the method of correlated vectors yields a modest positive or negative correlation between score gains and *g* loadings.

2. Method

2.1. Test

The GATB (United States Department of Labor, 1970; van der Flier & Boomsma-Suerink, 1994) is a test of general intelligence with eight subtests: *Three-Dimensional Space* measures Visualization (g_v), *Vocabulary* measures Induction (g_{fl}) and Lexical Knowledge (g_{cr}), *Arithmetic Reason* measures Quantitative Reasoning (g_{fl}), *Computation* measures Numerical Ability (g_{cr}), *Tool Matching* measures Perceptual Speed (g_s), *Form Matching* measures Spatial Relations (g_v), *Name Comparison* measures Perceptual Speed (g_s) and Numerical Ability (g_{cr}), and *Mark Making* measures Aiming (General Psychomotor Speed). There are also two additional tests for finger dexterity: *Assemble* and *Disassemble*; as well as two more tests for manual dexterity: *Place* and *Turn*.

2.2. Samples

Sample 1: Workers representative of the general working population from 1947. The first general working population norms for the GATB were based on 519 employed workers (U.S. GATB manual, 1970). It was recognized that the sample probably was not truly representative of the general working population, but since it

did include a wide range of occupational classifications, it was believed to yield a reasonably close approximation to test performance typical of the general working population. The date for the first sample is not explicitly given in the GATB manual, but close reading of the text suggests the year 1947. The mean age of the sample is 30.4 years (*SD* = 10.9 years) and the mean education is 11.0 years (*SD* = 2.4 years). This resulted in means and *SD*s for the GATB subtests, including the four tests of finger and manual dexterity.

Sample 2: Workers representative for the general working population from 1952. In 1952, general working population norms were established on the basis of a selected sample of 4,000 which was stratified to obtain proportional occupational representation of the general working population. The mean age of the sample is 30.4 years (*SD* = 9.9 years) and the mean education is 11.0 years (*SD* = 2.6 years). Means and S.D.s are reported for all twelve GATB subtests.

Samples 3+4: Dutch applicant bus drivers from 1975-1976 (sample 3), and 1983-1985 (sample 4). The Dutch GATB manual reports that for a specific time period a random sample (*N* = 110 for sample 3, and *N* = 1091 for sample 4) was taken from all persons who applied for positions of bus driver at regional bus companies and were tested at Dutch Railways selection centers.

Sample 5: Dutch applicant bus drivers from 1988-1992. For a specific time period, a random sample (*N* = 221) was taken from all persons who applied for positions of bus driver at regional bus companies and were tested at Dutch Railways selection centers (data are taken from te Nijenhuis, 1997).

Samples 6-8: Dutch 16-year-old students in higher general secondary education representative for the years 1975, 1985, and 2005, respectively. The Dutch GATB manual (van der Flier & Boomsma-Suerink, 1994, pp. 148-153) and an update of the manual (Akkerman, 2011) give large representative samples for various school types from 1975, 1985, and 2005. The sample sizes for students from higher general secondary education were *N* = 130 for 1975, *N* = 270 for 1985, and *N* = 498 for 2005.

2.3. Comparison of Samples

The mean scores of various cohorts were compared. 1) GATB scores from workers representative for the general working population from, respectively, 1947 and 1952 were compared. 2) Dutch GATB scores of three groups of applicants for the position of bus driver from 1975-1976, 1983-1985, and 1988-1992, respectively, were compared. 3) Dutch GATB scores of 16-year-old students in higher general secondary education in 1975, 1985, and 2005, respectively, were compared.

2.4. Statistical Analyses

2.4.1. Gain scores.

Standardized gain scores were computed by subtracting the score of the earlier sample on the test from the score of the later sample on the same test, and dividing the difference by the standard deviation of the earlier sample.

2.4.2. g loadings.

In general, *g* loadings were computed by submitting a correlation matrix to a principal axis factor analysis and using the loadings of the subtests on the first unrotated factor. In some cases *g* loadings were taken from studies where other procedures were followed; these procedures have been shown empirically to lead to highly comparable results.

2.4.3. Method of correlated vectors.

The method of correlated vectors requires the computation of a vector of *g* loadings and a vector of gain scores that are subsequently correlated. Pearson correlations between the standardized score gains and the *g* loadings were computed. The method of correlated vectors was only employed on samples in which the vast majority of subtests showed score gains.

3. Results

3.1. Gains

Tables 1-3 show the gains on the GATB for various samples. The general picture in Tables 1 and 2 is an increase while Table 3 shows both an increase and a decrease. The aggregated scores show a much clearer pattern than do the scores on the individual tests. The subtests closest to traditional classroom subjects are Vocabulary, Arithmetic Reason, and Computation; and the tests of Broad Visual Perception are Three Dimensional Space, Tool Matching, and Form Matching. Name Comparison has both a visual and a scholastic component and therefore does not fit unambiguously into one of the two previous clusters. Mark Making and the four dexterity tests have low *g*-loadings and low correlations with the other, more highly *g*-loaded tests. Table 4 shows the standardized gains per decade for the combination of visual tests and scholastic tests, respectively.

It appears that the gains on the visual cluster have remained roughly the same between 1917 and 1989. Further, between 1917 and about 1960 the gains on the scholastic cluster were roughly comparable, and then the scores started to drop dramatically. The pattern in Table 3 is very clear: between 1975 and 1985 there is a gain in the scores on visual tests and a large drop in the scores on scholastic tests; the differences between 1975 and 2005 show the effects are in the same direction and that they become much stronger. Flynn (2001) describes a similar shift in the US data around 1948, but the Dutch data suggest a shift around 1960 for the Netherlands. Flynn hypothesizes that this is because in the middle of the 20th century schools shifted their emphasis from the traditional classroom subjects of reading, writing, and arithmetic to reasoning with de-contextualized problems.

Table 1 shows a decrease in Mark Making, but an average increase for the four dexterity tests. Table 2 shows both a decrease of –0.03 *SD* per decade and an increase of 0.07 *SD* per decade on Mark Making. In sum, there is a modest increase on these subtests.

Table 1. Means and SDs on the General Aptitude Test Battery for the two General Working Population Samples From 1947 (N = 519) and 1952 (N = 4000), Standardized Score Gains, and g Loadings.

Test	1947		1952		*d*	*g*
	m	*SD*	*m*	*SD*		
Three-dimensional space	17.5	6.9	19.0	7.0	0.17	.64
Vocabulary	20.9	9.2	21.5	9.4	0.07	.68
Arithmetic Reason	8.7	3.4	9.9	3.8	0.35	.69
Computation	26.6	7.6	28.1	8.1	0.20	.72
Tool Matching	21.2	5.7	21.8	6.1	0.11	.70
Form Matching	26.3	7.8	26.9	8.1	0.08	.72
Name Comparison	71.3	20.1	70.7	22.1	-0.03	.79
Mark Making	71.0	9.7	69.5	10.3	-0.15	.67
Place	88.0	8.7	89.8	8.6	0.21	.40
Turn	101.0	8.7	100.8	9.6	-0.03	.41
Assemble	27.6	4.6	28.3	4.6	0.15	.40
Disassemble	28.8	3.7	29.5	3.7	0.19	.46

Note. *g* loadings were computed from the intercorrelations for the 1952 sample, as reported in the GATB manual (1970, p. 30); $N = 4000$.

3.2. Method of Correlated Vectors

We used the method of correlated vectors only on samples in which the vast majority of subtests showed score gains. Since the samples of students in higher general secondary education showed dramatic decreases on half the GATB subtests, they were excluded from the analyses (see above for explanation). For the general working population samples from 1947 and 1952 the score gains correlated $r = -.19$ with *g* loadings. For the applicant bus drivers from 1975-1976 and 1988-1992 the score gains correlated $r = .04$ with *g* loadings. For the Dutch applicant bus drivers from 1975-1976 and 1983-1985 the score gains correlated $r = .35$ with *g* loadings. For all three studies this results in an *N*-weighted $r = -.07$.

Table 2. Means and SDs on the Dutch General Aptitude Test Battery for Applicant bus drivers from 1975-1976, 1983-1985, and 1988-1992, Standardized Score Gains, and g Loadings

Test	1975-76		1983-85		1988-92		d_1	d_2	*g*
	m	*SD*	*m*	*SD*	*m*	*SD*			
Three-dimensional space	20.02	6.18	19.80	5.90	21.17	5.70	-0.04	0.19	.58
Vocabulary	20.01	6.33	20.32	6.77	24.50	7.18	0.05	0.71	.68
Arithmetic Reason	12.72	3.39	12.30	3.16	13.39	3.58	-0.12	0.20	.68
Computation	21.49	5.61	22.06	4.71	21.61	4.93	0.10	0.02	.67
Tool Matching	27.47	5.90	28.13	5.55	28.42	5.69	0.11	0.16	.49
Form Matching	25.16	5.71	26.68	5.85	28.45	5.98	0.27	0.58	.53
Name Comparison	54.81	12.83	55.39	12.01	60.35	13.87	0.05	0.43	.62
Mark Making	67.14	7.66	66.91	8.52	67.68	9.75	-0.03	0.07	.14

Note. 1975-1976: $N = 110$; 1983-1985: $N = 1091$; 1988-1992: $N = 212$.
d_1: gain from 1975-1976 to 1983-1985; d_2: gain from 1975-1976 to 1988-1992.
g loadings from de Wolff & Buiten (1963).

4. Discussion

Secular score gains are predominantly driven by environmental factors and might these factors also be responsible for group differences in intelligence? Is the pattern of secular score gains the same as the pattern of group differences? Group differences are strongly linked to *g* loadings, but the literature up to now yields conflicting findings concerning the link between *g* loadings and score gains, so additional studies are required. We used a number of datasets, which showed that the method of correlated vectors yields a small negative average correlation between score gains and *g* loadings. So, the general picture is now that of a small negative correlation between *g* loadings and gains. It appears that the Flynn effect and group differences have different correlations with *g* loadings, which suggests they have different causes.

Table 3. Means and SDs on the Dutch General Aptitude Test Battery for 16-year-old Students in Higher General Secondary Education in 1975, 1985, and 2005, Standardized Score Gains, and g Loadings

Test	1975		1985		2005		d_1	d_2	*g*
	m	*SD*	*m*	*SD*	*m*	*SD*			
Three-dimensional space	21.55	5.32	21.76	5.46	24.03	5.49	0.04	0.47	.58
Vocabulary	29.72	5.15	27.70	4.94	23.65	4.84	-0.39	-1.18	.68
Arithmetic Reason	14.92	3.05	14.29	3.27	11.42	3.02	-0.21	-1.15	.68
Computation	25.11	4.64	22.21	4.04	19.50	4.18	-0.63	-1.21	.67
Tool Matching	33.56	5.04	34.44	5.39	36.14	5.67	0.18	0.51	.49
Form Matching	30.28	5.79	30.90	6.51	32.37	6.54	0.11	0.36	.53
Name Comparison	30.28	5.79	30.90	6.51	32.37	6.54	0.11	0.36	.53
Mark Making	75.24	11.82	66.06	10.48	65.02	11.56	-0.78	-0.87	.62

Note. 1975: $N = 130$; 1985: $N = 270$; 2005: $N = 498$.
d_1: gain from 1975 to 1985; d_2: gain from 1975 to 2005.
g loadings from de Wolff & Buiten (1963).

As expected, we found overall gains on the large majority of measures. Previous studies of the Flynn effect made extensive use of the Wechsler tests and the various versions of Raven's Progressive Matrices. This study has now also shown clear Flynn effects for the General Aptitude Test Battery. Flynn (2006) states that there are few datasets on adults, and most of these are on military conscripts. We added data from samples of applicant bus drivers and workers representative for the general working population and found clear secular score gains.

Table 4. Standardized Gains per Decade (in SDs) in GATB Scores from Tables 1-3.

Born	Sample	Gain per decade	
		Visual	Scholastic
1917-1922	Working population	0.24	0.42
1945-1954	Applicant bus drivers	0.13	0.01
1945-1960	Applicant bus drivers	0.21	0.21
1959-1969	16-year-old students	0.11	-0.41
1959-1989	16-year-old students	0.15	-0.39

Note. Visual subtests: Three Dimensional Space, Tool Matching, and Form Matching; Scholastic subtests: Vocabulary, Arithmetic Reason, and Computation.
The average age for the applicant bus drivers from 1988-1992 was 29.91 years (data from te Nijenhuis, 1997) and we take this value as the average of the other two samples of applicant bus drivers.

Secular gains of about one *SD* have been shown for infants on tests for motor development, such as the Griffiths scale and the Bayley scales (see Lynn, 2009). We found Flynn effects for adults on perceptual motor tests, such as the four GATB tests which measure Finger and Manual dexterity as well as the GATB tests Mark Making, which measures Aiming. So, secular increases in perceptual motor ability and attention occur not only for toddlers and young children, but also for adults. The gains on motor tests seem to last into adulthood.

The method of correlated vectors yielded small to modest positive and negative correlations between score gains and *g* loadings in all cases where there were Flynn effects on the large majority of subtests, with an *N*-weighted r = -.07. The combined literature is now suggestive of a modest negative relationship between *g* and *d*.

5. Conclusion

The research literature overwhelmingly showed that group differences and *g* loadings are strongly correlated, but the literature on the Flynn effect and *g* loadings up to now showed conflicting findings. However, all three additional studies in the present study taken together show a small negative correlation between *g* loadings and secular gains. The strong positive correlation of Flynn (2000) and Nisbett (2009) appears to be an outlier. There are strong differences of opinion about the meaning of these findings, so more research is clearly needed.

5.1. Thoughts About the Future of the Area

The nomological net of group differences, secular score gains, and *g* loadings needs to be explored more fully, using the method of correlated vectors. For instance, Jensen (1998) showed that variables such as head size, brain volume, brain's gray matter, brain's evoked potential, brain glucose metabolic rate, peripheral nerve conduction velocity, brain pH, body symmetry, inbreeding depression, and hybrid vigor yield high positive correlations with *g* scores. Moreover, the gains resulting from test-retest and test training on IQ tests correlate perfectly negatively with *g* scores (te Nijenhuis, Voskuijl, & Schijve, 2001; te Nijenhuis, van Vianen, & van der Flier, 2007). For instance, various causes have been hypothesized for the Flynn effect, including schooling, nutrition, health care, heterosis, GDP, urbanization, smaller families, increased exposure to movies and optical displays, the dissemination of visual-spatial toys, increased environmental complexity, and teacher to student ratio. Usable data can be tested with the method of correlated vectors, to see whether they have a pattern that is more comparable to the pattern of secular score gains or to the pattern of group differences.

The method of correlated vectors is not a strong statistic; it's simply the correlation between a small number of observations. However, it is a golden combination with one of the strongest methods available in science, namely psychometric meta-analysis (Hunter & Schmidt, 2004), which estimates what the results of studies would have been if all studies had been conducted without methodological limitations or flaws. What is badly needed is a psychometric meta-analysis of all the studies correlating *g* loadings and secular score gains. The hypothesis that the large variety in effect sizes can be fully explained by statistical artifacts can be tested in a psychometric meta-analysis.

References

Akkerman, A.E. (2011). *GATB Handleiding* [GATB Manual]. The Netherlands, Utrecht: Stichting G.A.T.B.- Research.

Colom, R., Juan-Espinosa, M., & García, L. F. (2001). The secular increase in test scores is a "Jensen effect". *Personality and Individual Differences, 30*, 553-559.

de Wolff, C. J., & Buiten, B. (1963). Een factoranalyse van vier testbatterijen [A factor analysis of four test batteries]. *Nederlands Tijdschrift voor de Psychologie, 18*, 220-239.

Dickens, W.T., & Flynn, J.R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 198*, 346-369.

Dickens, W.T., & Flynn, J.R. (2006a) Black Americans reduce the racial IQ gap: Evidence from standardization samples. *Psychological Science, 17*, 913-920.

Dickens, W.T., & Flynn, J.R. (2006b). Common ground and difference. *Psychological Science, 17*, 923-924.

Flynn, J. R. (1984). The mean of IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51.

Flynn, J. R. (1999). Evidence against Rushton: The genetic loading of WISC-R subtests and the causes of between-group IQ differences. *Personality and Individual Differences, 26*, 373-379.

Flynn, J. R. (2000). IQ gains, WISC subtests and fluid *g*: *g* theory and the relevance of Spearman's hypothesis to race. In G. R. B. J. Goode (Ed.), *The nature of intelligence* (pp. 202-227). New York: Wiley.

Flynn, J. R. (2001). *The history of the American mind in the 20th century*. Paper presented at Spearman Seminar, Sydney, Australia.

Flynn, J. R. (2006). Efeito Flynn: Repensando a inteligência e seus efeitos [The Flynn Effect:
Rethinking intelligence and what effects it]. In C. Flores-Mendoza & R. Colom (Ed.), *Introdução à Psicologia das Diferenças Individuais* (pp. 387-411). Brazil, Porto Alegre: ArtMed.

Flynn, J.R. (2007). *What is intelligence?* Cambridge University Press.
Flynn, J.R. (2010). The spectacles through which I see the race and IQ debate. *Intelligence, 38*, 363-366.
Hunter, J.E., & Schmidt, F.L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings, second edition*. London: Sage.
Jensen, A. R. (1987). Individual differences in mental ability. In J. A. Glover & R. R. Ronning (Eds.), *Historical foundations of educational psychology*. New York: Plenum.
Jensen, A. R. (1998). *The g factor: The science of mental ability*. London: Praeger.
Lakatos, I, & Musgrave, A. (1974). *Criticism and the growth of knowledge*. Cambridge University Press.
Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the Development Quotients of infants. *Intelligence, 37*, 16-24.
Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations*. London: Praeger.
Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other trends. *Psychological Review, 114*, 806-829.
Must, O., Must, A., Raudik, V. (2003). The secular rise in IQs: In Estonia the Flynn effect is not a Jensen effect. *Intelligence, 167*, 1-11.
Nisbett, R. E. (2009). *Intelligence and how to get it*. New York: Norton.
Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3*, 346-353.
Rijsdijk, F. V., Vernon, P. A., & Boomsma, D. I. (2002). Application of hierarchical genetic models to Raven and WAIS subtests: A Dutch twin study. *Behavior Genetics, 32*, 199-210.
Rushton, J. P. (1989). Japanese inbreeding depression scores: Predictors of cognitive differences between blacks and whites. *Intelligence, 13*, 43-51.

Rushton, J. P. (1999). Secular gains in IQ not related to the *g* factor and inbreeding depression - unlike Black-White differences: A reply to Flynn. *Personality and Individual Differences, 26*, 381-389.
Rushton, J. P., Bons, T. A., Vernon, P. A., & Čvorovič, J. (2007). Genetic and environmental contributions to population group differences on the Raven's Progressive Matrices estimated from twins reared together and apart. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *274*, 1773–1777.
Rushton, J. P., Čvorovič , J., & Bons, T. A. (2007). General mental ability in South Asians: Data from three Roma (Gypsy) communities in Serbia. *Intelligence, 35*, 1-12.
Rushton, J. P., & Jensen, A. R. (2003). African-White IQ differences from Zimbabwe on the Wechsler Intelligence Scale for Children-Revised are mainly on the *g* factor. *Personality and Individual Differences, 2003*, 177-183.
Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*, 235-294.
Rushton, J.P., & Jensen, A.R. (2010). Editorial. The rise and fall of the Flynn effect as a reason to expect a narrowing of the Black–White IQ gap. *Intelligence*, *38*, 213-219.
Rushton, J. P., Skuy, M., & Bons, T. A. (2004). Construct validity of Raven's Advanced Progressive Matrices for African and non-African engineering students in South Africa. *International Journal of Selection and Assessment, 12*, 220-229.
Rushton, J. P., Skuy, M., & Fridjohn, P. (2002). Jensen Effects among African, Indian, and White engineering students in South Africa on Raven's Standard Progressive Matrices. *Intelligence, 30*, 409-423.
Rushton, J. P., Skuy, M., & Fridjohn, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White engineering students in South Africa. *Intelligence, 31*, 123-137.

Spitz, H. H. (1988). Wechsler subtest patterns of mentally retarded groups: Relationship to *g* and to estimates of heritability. *Intelligence, 12*, 279-297.

te Nijenhuis, J. (1997). *Comparability of test scores for immigrants and majority group members in the Netherlands.* Unpublished Unpublished doctoral dissertation, Vrije Universiteit, Amsterdam.

te Nijenhuis, J., de Jong, M.-J., Evers, A., & van der Flier, H. (2004). Are cognitive differences between immigrant and majority groups diminishing? *European Journal of Personality*, 405-434.

te Nijenhuis, J., & van der Flier, H. (2003). Immigrant-majority group differences in cognitive performance: Jensen effects, cultural effects, or both? *Intelligence, 31*, 443-459.

te Nijenhuis, J., van Vianen, A. E. M., & van der Flier, H. (2007). Score gains on *g*-loaded tests: No *g*. *Intelligence*, *35*, 283-300.

te Nijenhuis, J., Voskuijl, O. F., & Schijve, N. B. (2001). Practice and coaching on IQ tests: Quite a lot of g. *International Journal of Selection and Assessment, 9*, 302-308.

U.S. Department of Labor (1970). *Manual for the USTES General Aptitude Test Battery. Section III. Development*. Washington, DC: United States Department of Labor.

van der Flier, H., & Boomsma-Suerink, J. L. (1994). *Handboek GATB [GATB manual]*. The Netherlands, Amsterdam: Stichting G.A.T.B.-Research.

Wicherts, J. W., Dolan, C. V., Oosterveld, P., van Baal, G. C. V., Boomsma, D. I., & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence, 32*, 509-537.

Woodley, M. A. (2011). Heterosis doesn't cause the Flynn effect: A critical examination of Mingroni (2007). *Psychological Review*, 118, 689-693.

Chapter 4

African Cognitive Ability: Research, Results, Divergences and Recommendations

Heiner Rindermann
Department of Psychology, Chemnitz University of Technology
Wilhelm-Raabe-Str. 43, D-09107 Chemnitz, Germany

Abstract

In the past different researchers have come to diverging cognitive ability estimates for people in Africa and of African descent. This paper checks the validity of past results by comparing them with outcomes of two new psychometric test studies from East and South Africa; with results from student assessment studies; with predictions based on those variables which, outside Africa, correlate most strongly with intelligence; and by comparing them with further indicators of cognitive ability (descriptions of everyday life and human accomplishment). Integrating these cognitive ability measures with the application of several corrections (due to the higher age of students in Africa, lower African school enrollment, selectivity of samples and higher African secular IQ rise), the best estimate for an African average is IQ 75. Finally, possible environmental and genetic (evolutionary, therefore past environmental) causes are discussed and suggestions are given how to enhance cognitive development in African countries.

1. Cognitive ability research in Africa

Cognitive abilities and differences between groups have been a hotly debated subject in research, with the differences between "Whites" (people of European descent) and "Blacks" (people of sub-Saharan African descent), in particular, causing scientific and non-scientific conflicts (e.g., Segerstråle, 2000). Within the US and other Western countries the mean cognitive ability difference in various tests (from psychometric IQ to student competence tests as SAT or TIMSS) averages around one standard deviation (Gonzales, 2000, p. 59, 61; Hunt, 2011, p. 411f.). The ability levels for Africans in Africa are the subject of strong disagreement. Rushton studied positively selected samples (university engineering students in South Africa; Rushton, Skuy & Fridjhon, 2003), but the mean differences between Africans and Europeans (14 IQ points) were similar to the ones found in Western countries. Lynn and Vanhanen (2006) estimated that sub-Saharan African countries had a mean IQ of 70. Wicherts, Dolan and Maas (2010) used a different selection procedure and came to a mean IQ of 82.

This paper presents data from two independently conducted psychometric intelligence studies, reanalyze student achievement data, use predictive variables, which have been validated in the rest of the world, to estimate mean African ability, and give a short overview on everyday life indicators of cognitive abilities.

2. Cognitive ability

Cognitive ability (equivalent to cognitive competence) refers to the ability to think (intelligence), knowledge (true and relevant knowledge) and the intelligent use of this knowledge. A broad concept of "intelligence" also includes knowledge aspects ("crystallized intelligence"). Cognitive ability enhances the individual's understanding of concepts and causal relationship, it increases insight, foresight, and rationality. It leads to proximal consequences, such as higher quality of work and more reasonable decisions in everyday life. Higher cognitive ability also improves

individuals' access to better environments and enables individuals, institutions, societies, and cultures to improve the quality of the available environment. Cognitive ability also brings about distal consequences, such as greater wealth and health; a more democratic society; political and economic liberty; a more complex culture; and longitudinally, by backward effects of these environmental factors, enhanced intelligence (e.g., Rindermann, 2012; Rindermann & Meisenberg, 2009; Rindermann & Thompson, 2011).

3. Preliminary remarks on research

Science sometimes creates tensions between research findings and society. Epistemic-scientific principles can be at conflict with legitimate economic, cultural or ideological interests, usually represented by the political class, media, church, intellectuals or the public. However, even in hotly debated areas of research, fundamental principles of scientific thinking have to be applied. Science is a process based on epistemic rationality guided by logicality, empiricity and argumentativity. Scientists write for an abstract, rational reader who can be convinced (an ability and a willingness) through argumentation using logic, empirical facts and systematic reasoning. Freedom of research and respect for others in their scientific endeavor will help the entire scientific community to progress (Ceci & Williams, 2009; Flynn, 2007).

Other, in their fields legitimate orientations are empirically relevant, but not for science as an endeavor to pursue the truth. In science, from an epistemic-scientific view, only the truth or falseness of statements matter and an angel's truth is as true as a devil's truth. It is irrelevant, if a statement is blue or red, progressive or conservative, up or down, welcomed by the *x* or *y*, right or left, pc or non-pc, published here or there, welcomed and repeated by the right or wrong people. Of importance is, if it is correctly describing the world and explaining it, and secondly, if it is new and develops stimulating theoretical approaches.

Not all those arguing about intelligence have observed such rules, and participants of past conflicts have suffered from offensive treatment including violent attacks (Gottfredson, 2010; Nyborg, 2003). However, intellectual conflicts are not new in the history of thought, as the fate of scholars like Thomas Aquinas, Galilei, Spinoza, and Darwin shows. From a contemporary perspective many past disputes sound quite ridiculous and their formerly not questionable "arguments" are today scientifically and ethically disapproved. Even then, the conflicts have been important in developing in the long term a climate of argumentation and thinking. The frequently difficult process of Enlightenment is not strengthened if people shy away from such conflicts.

4. Cognitive ability measures and samples – method

A detailed description of data and method is found in the *Supplementary data file*. Briefly:

1. Data from *N*=174 students in grades 8 and 9 were collected in *Tanzania* in 1999 and 2000 by the use of three culture-reduced intelligence tests (APM, MRT, LPS). For interpretation, "FLynn" corrections of German norms, Greenwich IQ correction, school enrollment correction and the larger FLynn effect in Africa were considered (we assume for the last two decades in Africa a stronger secular rise of intelligence than in western countries; Daley et al., 2003).
2. The cognitive ability of blind, visually handicapped and non-blind students in the age of 10 to 16 years in *South Africa* (*N*=153, African ancestry 41%, Mixed/"Coloured" 34%, European 25%) and *Austria* (*N*=63) were compared by the use of WISC-IV working memory and verbal comprehension scales. The results from 2008 were corrected for selection bias in South Africa, composition of the South African population, lower school enrollment in older ages and the FLynn effect. In this study the visual handicap itself is not important, but the difference between people of

different descent and the possible furtherance effect by a visual handicap on working memory.

3. Results from student assessment studies (1964-2009, in Africa mainly TIMSS, PIRLS and SACMEQ; N_c=14 countries) collected by four different research groups were combined, corrected (for school enrollment, age and grades) and restandardized to international Greenwich norm (UK set at IQ 100).[1]

4. Finally, in a regression analysis the two variables most highly predictive of cognitive ability (in the rest of the world outside sub-Saharan Africa), but theoretically distinct, were selected to predict cognitive ability levels in countries populated by a majority of people with African ancestry (N_c=52 sub-Saharan-African and Caribbean countries). To have a predictive value a causal theory is not necessary. The predictors stem from the two competing paradigms, nurture vs. nature: The first is the *Human Development Index* (HDI; with IQ r=.75 in N_c=107 non-African countries), the second is *skin brightness* (or skin reflectance with IQ; r=.75, N_c=82; correlating with each other: r=.70, N_c=78). Their results (SD corrected) were averaged. HDI correlates more strongly with cognitive ability than education or wealth (r=.67, .53/.60, last logged wealth). Skin brightness cannot have a direct causal effect on cognitive ability. This variable and its relation to evolutionary development may be distressing normatively. It is also related to history, slavery, apartheid and other physical and psychological maltreatment of African people. The same ambiguity is true for HDI: Although it is a clear environmental indicator, it depends on cognitive ability and differences can depend on yet unknown genetic factors.

---

[1] Postscript 2013: There are two further local African studies which I did not include (when writing the paper simply not known): Monitoring Learning Achievement (MLA) and Programme d'Analyse des Systèmes Éducatifs (PASEC). I will add them in further international studies. However, they do not use internationally anchored scales (only results for African countries). Therefore, they only help in within African comparisons.

As a reference point Greenwich-IQ was used. Results were compared with psychometric IQ measures from Lynn and Vanhanen (2006, updated), psychometric IQ measures from Wicherts et al. (2010), and a cognitive ability measure consisting of student assessment and psychometric IQ test results. Finally a difference was calculated between the here newly predicted and formerly published ability levels.

5. Results[2]

Table 1. IQ values for countries with African majority.

Country	S1-2	SAS-k	Lynn 79	Wich	Rind	HDI-IQ	SB-IQ	HS-IQ	Diff
Angola			(68)		(64)	69	68	68	4
Bahamas			(84)		(80)	94		94	14
Barbados			80		80	94		94	14
Benin			(70)		(66)	71	68	69	3
Botswana		76	71		76	84	73	78	3
B. Faso			(68)		(64)	62	65	64	0
Burundi			(69)		(65)	61	68	64	-1
Cameroon			64		64	72	64	68	4
C. Verde			(76)		(72)	77		77	5
C. Afr. R.			64	71	64	63	66	64	1
Chad			(68)		(64)	61	65	63	-1
Comoros			(77)		(73)	70		70	-3
Congo (Br.)			64	78	64	74	70	72	8
C. d'Ivoire			71		71	68	71	70	-1
Djibouti			(68)		(64)	69	73	71	7
Dominica			67		67				
Eq. Guinea			(64)		(60)	78	73	75	16
Eritrea			76	85	76		71	71	-4
Ethiopia			69	69	69	64	70	67	-2
Gabon			(64)		(60)	85	68	76	17
Gambia			64		64	68	62	64	1
Ghana		69	71	73	65	73	68	70	6
G-Bissau			(67)		(63)	61	66	64	1
Guinea			67		67	64	64	64	-3

[2] Details can be found in the supplement at the end of the paper.

Country	S1-2	SAS-k	Lynn 79	Wich	Rind	HDI-IQ	SB-IQ	HS-IQ	Diff
Haiti			(67)		(63)	69		69	6
Jamaica			71		71	88		88	17
Kenya		81	72		75	73	69	71	-4
Lesotho		65	(67)	80	64	70	68	69	5
Liberia			(67)		(63)	62	67	64	2
Madagascar			82	82	82	71		71	-11
Malawi		61	60		60	67	66	67	7
Mali			68	74	68	62	72	67	-1
Mauritania			(76)		(72)	71	79	75	3
Mozambique		73	64		74	61	63	62	-12
Namibia		65	72	74	69	82	67	75	5
Niger			(69)		(65)	59	68	64	-1
Nigeria		76	69	84	76	70	67	68	-8
Rwanda			76		76	67	68	68	-8
Sao T. & P.			(67)		(63)	74		74	11
Senegal			71	66	71	69	64	67	-4
S. Leone			64	91	64	63	68	65	1
Somalia			(68)		(64)		68	68	4
S. Africa	85	68*	72	77	68	82	75	78	10
St. Lucia			62		62				
Sudan			77		77	67	71	69	-8
Swaziland		77	(68)		68	75	70	72	4
Tanzania	73	77	72		74	68	66	67	-6
Togo			(70)		(66)	70	68	69	3
Uganda		70	73	84	72	70	64	67	-5
Zaire (C., K.)			65	76	65	58	68	63	-2
Zambia		62	71	79	68	68	64	66	-2
Zimbabwe		76	72	82	75	51	68	59	-16

Notes: Countries with African majority populations. *S1-2*: Studies 1 and 2 using psychometric tests, reference year 2010, samples probably positively selected; *SAS-k*: student assessment studies from 1964-2009 combined, standardized and corrected and non-corrected averaged; *Lynn79*: measured (and in parentheses estimated) psychometric intelligence test data by Lynn & Vanhanen (updated 2010; reference year 1979); *Wich*: measured psychometric intelligence test data by Wicherts et al., UK norm; *Rind*: measured (and in parentheses estimated) psychometric intelligence (reference year 1979) and student assessment test data by Rindermann; *HDI-IQ*: IQ predicted by the use of the HDI; *SB-IQ*: IQ predicted by the use of skin brightness; *HS-IQ*: IQ predicted by the use of HDI and skin brightness (mean); *Diff*: difference between measured (sometimes estimated) cognitive ability (Rind) and by HDI and skin brightness together predicted IQ.

The mean of the two newer unpublished psychometric studies (N_c=2 countries, Table 1, column 1, S1-2) in 2010 norms is IQ 79. The mean result in student assessment studies (N_c=14 countries, Table 1, column 2, SAS-k) in international norms (majority of studies 1995-2007) is IQ 71. Lynn's 1979 norm data have a mean of IQ 70 (N_c=52), or of IQ 76 (in 2010) after FLynn correction (larger than in UK since 1979). Wicherts' collection (N_c=17) results in IQ 77. Rindermann's collection results in IQ 68 (N_c=52), or IQ 73 after FLynn correction. HDI predicts a mean IQ of 70 (N_c=48), skin brightness IQ 68 (N_c=42), both together IQ 70 (N_c=50). FLynn-corrected for 2010, the predicted IQs are around IQ 75. The total mean range of all studies and different values discussed by the authors is between 68 and 82. If one considers only 2010 estimates and excludes outliers the range is between 71 and 78. Student assessment studies with their larger school-related test content and therefore larger dependence on educational quality seem to boost the difference to more developed countries.

HDI-based IQ predictions are higher than measured IQs for countries in the Caribbean and in South-Africa, and they are higher for countries with formerly only by neighboring countries' test results estimated data. In the Caribbean and in South-Africa the general living conditions are better than expected by cognitive ability. The near Western world may positively affect the living conditions, and the additional effects of minorities and of enduring institutions of past mother countries (UK, Netherlands) could influence the fate of nations even today (Harrison, 2006). Past estimations (especially with downward corrections) seem to underrate IQs. Other countries such as Zimbabwe and Mozambique are less developed than expected by their measured cognitive ability level, perhaps because of past civil war or bad government. "Skin brightness", compared to "HDI", is the better predictor (less deviation of the predicted from measured and estimated values: mean squared difference for the HDI-predictions is D^2=76.14, for the skin brightness-predictions D^2=31.09). In the same sample of N_c=17 countries the Lynn-data correlate more highly with the predicted IQ than the Wicherts-data (r=.34 vs. r=-.02). The

deviance from predicted IQ is smaller for the Lynn-IQ (D^2=29.86 vs. 149.50).

The described IQ means vary between IQ 68 and 78. Averaging the given means for 2010 results in an estimated IQ of around 75 for African majority countries. This result is based on many assumptions and corrections, such as properly given tests and understood test instructions, sample representativity and school enrollment rates and their corrections, mean IQ growth per year at school and per age year, and a higher African FLynn-effect. Researchers diverge in selection and correction criteria. Given the quality of the data, it is *not possible* to come to a really *precise result*. Compared to other developing regions the results are not astonishing (e.g. Bali in South-East Asia with IQ 79; Rindermann & te Nijenhuis, 2012). There are further arguments indicating overestimation (student assessment studies not corrected for older age) or underestimation (less test-wiseness). The usual phrase "further research is needed" is very appropriate here: We need representative samples of the ages 10 to 70; samples representative of the full range of school education, including the share without or with only little education; the use of fluid (school-distant) and crystallized (school-near) cognitive ability tests; and up-to-date norms from Great Britain. Furthermore, as African samples have less cognitive task experience, the estimates could be increased by a short test training or a more general cognitive training (Skuy et al., 2002).

The South African study compared people of African and European descent from South Africa and Austria and from different SES levels. In a regression analysis the most important predictor (highest β) was evolutionary-ethnic background, then SES and then country. Additionally, the more visually handicapped a person was, the stronger was his or her working memory (IQ 119, 108 and 105), perhaps because visually handicapped persons develop a stronger working memory to compensate for their reduced ability to store information in written form. This is strong support for the *malleability* of even a very basic cognitive ability!

6. Further indicators

Test studies of intelligence differences across nations have been criticized for artificiality of test situation, low representativity and validity, and cultural bias. Some claimed that results stand for "How well can *they* do *our* tricks?" and not for cognitive ability. The concept of intelligence and tests were developed in the Western world as well as e.g. gravity and the meter. But taking e.g. instead of height the weight or using non-metric units does not change people's height. It is important to properly define a construct and find appropriate ways to measure it. Then it is unimportant where and by whom it was developed. However, it is true that research should look for further indicators of cognitive ability, especially in everyday life and in cultural sediments. Before reading the following descriptions, we would welcome it if readers think for five minutes which indicators they would choose to estimate cognitive ability.

Researchers such as Baker (1974) and Hart (2007) have tried to develop indicators of intelligence in everyday life and cultural artifacts. Examples of such indicators include invention and use of script, of the wheel for transport, pottery, the domestication of animals, the development of law, use of abstract numbers, appreciation of knowledge and school, of ethical systems, no torture and self-mutilation, hygiene, and the quality of architecture. Using these systematic comparisons across many cultures both authors estimated the cognitive ability level in Africa and of Africans as not being high.

Some may perceive such studies as biased. Every single indicator stands not only for cognitive ability but also for other phenomena depending on various conditions. If persons remain skeptical it is recommendable to develop better justified criteria and then look at the empirical material. For instance, it would be possible to assess the quality of art, such as the ability to draw and model realistic and dynamic pictures and figures; or to develop criteria for understanding nature and life (e.g., medical knowledge and effective treatment, astronomy). Benchmarks could be works

of ancient Greece or Nineveh, e.g. King Ashurbanipal hunting lions. Even in prehistoric times there may be found valuable works (e.g., Paleolithic cave painting; astronomic knowledge as Nebra sky disk and Stonehenge; the Maya calendar; inventions as yoke, compass, wheelbarrow, gunpowder and rice-farming). Of course, also African exceptional achievement will be found (e.g., Nigerian Nok culture, Ethiopian Lalibela, Great Zimbabwe). As in tests, A major critique of the outlined indicators is that they cover only past history. All peoples were in prehistory less developed than today. *Contemporary* information is necessary. Evidence could be provided by indicators of present-day *rationality* and belief systems. According to Piaget (1953) magic thinking is indicative of *preoperational thinking* (usually in Western samples ending at age 6, at least at age 11, corresponding to an adults IQ of 50-70; Rindermann, 2011). Many researchers, anthropologists and journalists (e.g., Caldwell, 2002; Dagona, 1994; Kabou, 1991; Signer, 2004) have documented sorcery, Voodoo, and fetishism in today's African populations. Similar historical reports on belief in witchcraft can be found in Europe until 250 years ago (and relics of magic thinking remain until today as a kind of not seriously taken accompaniment in daily life, e.g., astrology and homeopathy) indicating that *cognitive modernization* is a ongoing historical process. Another example is the widespread belief that AIDS is caused by supernatural powers (Caldwell, 2002; Oesterdiekhoff & Rindermann, 2007). Such preoperational thinking in Piaget's terms unfortunately includes members of the political elite, influencing politics and culture.[3] [4]

A further evidence is *exceptional achievement*. Murray (2003, p. 284) found using Science Nobel Prizes (1901-2000) as indicators of human accomplishment in Africa 0%. However, Nobel Prizes do not only depend exclusively on individuals'

[3] For instance, the president of Gambia, Yahya Jammeh, claimed that he is a healer and can cure AIDS within three days by the laying on of his hands.

[4] South African's president Jacob Zuma mentioned that taking a shower after sexual intercourse with a HIV-infected woman protected him against HIV transmission.

cognitive ability: Higher society ability levels lead to better schools and universities, a more stimulating cultural and social environment, nets of contacts to institutions which award prizes, and higher wealth attracting scientists and enabling the donation of awards. Individuals with exceptional abilities surrounded by non-supportive conditions may fail to show their potential. Cognitive ability differences are boosted and stabilized through their institutional effects ending in reinforcing feedback-loops. Eventually, relatively small mean differences lead to large proportional differences at the highest levels.

As last indicator *government effectiveness* achieves in sub-Saharan Africa (Rindermann, Sailer & Thompson, 2009), standardized like IQ (UK M=100, international SD=15), a mean score of 58 (N_c=48 countries). Individual attributes are translated into institutional-social ones having retroactive effects on individuals again.

7. Causes

Generally, within psychology two causal paradigms are distinguished: The *environmental* and the *genetic*. In educational, sociological or economic research based on intervention studies the environmental perspective dominates. Sometimes, especially at the group level, it is not unusual to stress one's opposition to the genetic paradigm to acquire credibility and more acceptance. E. g.: "No one at the symposium believed that culture is genetically determined; rather, all believed that culture is acquired." (Harrison, 2006, p. XIII)[5] Usually, the genetic paradigm is simply ignored. Within intelligence research, based on behavioral genetic studies, the hereditary view explaining individual differences dominates.

[5] Dawkins (2008/1992, p. 26) mentioned a "dogmatic and hysterical opposition to the very possibility of genetic variation in human mental abilities." Pinker (2007) judges the hypothesis of genetically based group differences in ability as the currently most "dangerous idea".

Undoubtedly, there are *environmental* conditions impairing in Africa cognitive development including knowledge acquisition: problems with nutrition and health care; parasite load; lower educational background of parents, peers and society; short, interrupted and lower quality school enrollment; large classes; poor school equipment; lower educational level of teachers; teacher absence; lower quality classes, schools and universities; rote learning; school fees; insufficient or too expensive transport to school; poverty forcing children to work and not to learn; corruption reducing investment in education; war and violence including against schools and teachers; irrationality in everyday life; and problems in government effectiveness (e.g., Glewwe & Kremer, 2006).

However, genetic theories are not excluded by showing strong environmental effects – and the environment is not excluded by assuming genetic effects. Firstly, genes also influence environment and culture – the "extended phenotype" (Dawkins, 2008/1992), but we do not know to what degree and how. Secondly, evolutionary genetic theories are in the long run environmental theories (e.g., Rushton, 2000/1995, 2004; Rushton & Rushton, 2004). As adaptationist views longitudinally contain the environment in genes via selection of the successfully adapted model, only a non-adaptationist view of evolution would lead to pure genetic effects. Thirdly, the smaller the difference between mean IQs of Africans in sub-Saharan countries and in Western countries (e.g., Wicherts et al., 2010, in their own summary IQ 82; in US in Greenwich-norm 84) the less important can be (poor) environmental conditions for explaining African-European IQ differences. Following Cavalli-Sforza (1997, p. 7724) African-Americans have on average 30% of their gene pool from Europeans. If genetic factors contribute to the African-European intelligence-difference, then this admixture would be also relevant for the higher African-American IQ compared to the one in Africa. With our lower African estimate of IQ 75 (vs. African-American IQ of 84) environmental factors causing lower African IQ outside

Western countries are highly compatible. Environmentalism is supported by larger test gaps!

Evolutionary theories never denied environmental factors, but stressed that genetic factors play a crucial role (e.g., Rushton & Jensen, 2005). There is no genetic determinism, but genetic influence. There is no direct impact of genes on behavior, but indirectly through coding enzymes and influencing neurological system. However, the possible causal chain, e.g. through brain size or mental speed, is far from being clear. Both proposed mediators explain only about d=0.19 (equivalent to 3 IQ; Hunt, 2011, p. 433f.).

What evidence supports the contribution of genetic factors? Although there are large overviews (e.g., Rushton & Jensen, 2005), we want to consider three further scientific arguments and two "social-scientific" and "political" "arguments":

1. The stability of comparatively lower cognitive ability levels – across countries, indicators, time and societies – denotes an effect of a *stable* and *strong determinant*. This can be culture; however, is culture as stable as genes and does culture exclude genetic effects?
2. In the South-African study the evolutionary-ethnic background was a more important predictor for cognitive ability than SES. The evolutionary-ethnic background does not only reflect differences in socioeconomic status. It could stand for genetic or not by SES covered environmental and cultural differences.
3. Outside sub-Saharan Africa national cognitive ability levels are highly related to haplogroup distributions, even after controlling for general development of nations (HDI; Rindermann, Woodley & Stratford, 2012).
4. Some believe that there is a public agreement that genes are not appropriate for explaining ability differences between groups. A closer look reveals within science the contrary view: In an older opinion poll among N=1020 experts (Snyderman & Rothman, 1987) 15% believed that only environment is relevant for Black-White IQ-differences, but 45% believed that *environment and genes* are relevant (1% only genes, rest no opinion). Of course, majority opinion is no criterion for truth. Furthermore, a recently

published textbook by a researcher well known for his lack of enthusiasm for genetic explanations of group differences stressed the possibility of genetic factors: "Rushton and Jensen (and Lynn) are correct in saying that the 100% environmental hypothesis cannot be maintained. Nisbett's extreme statement [genes play no role at all] has virtually no chance of being true." (Hunt, 2011, p. 434)[6] Similarly Wicherts et al. (2010, p. 17) do not exclude genes among possible causes: "Although it cannot be precluded that genetic effects play a role in the low IQ performance of Africans, we view environmental circumstances as potentially more relevant to the present-day difference in mean."

5. If we consider only the environment to be responsible for intelligence differences this has a positive impact for improving cognitive ability because we can change environment. Environment is a mainly human-made environment, made by us, by our society, our ancestors and our culture. However, this theory also blames low IQ people: Why did they not change the environment now and in the past? Denying any genetic effect implies putting all the responsibility on human factors and on Africans (if not given the responsibility for the low IQ of Africans to external factors). But the higher mean IQ and general living conditions (as life expectancy) of Africans in the US and in similar countries compared to Africans in Africa underscores the positive impact of European culture. A genetic explanation is firstly a scientific hypothesis, true or not, and has to be checked in this regard; only secondly is it a political or moral issue. In this second meaning a genetic explanation excuses in the positive (no blame and guilt) and negative (does not stimulate further effort) way. It is not accusing or disparaging, By contrast, it could motivate to put more effort into compensation (Lindsay, 2010).

---

[6] Nisbett disputes any genetic cause of group differences. However, in his 2009 book he argued for a genetic effect in explaining low spatial reasoning of Jews: "The single ethnic difference that I believe is likely to have a genetic basis is the relative Jewish incapacity for spatial reasoning." (Nisbett, 2009, p. 254)

8. Limitations, future and suggestions

Mean differences do not deny overlaps and large individual differences. There are significant overlaps between nations, and between biologically or culturally defined human subgroups. There are smart Africans and dull Europeans. We are strictly opposed to any political misuse of this research. From an ethical viewpoint, persons have to be treated as individuals and not as a mere representative of any group. E.g. if job applicants have to be selected, only individual achievement is allowed to count (no racial discrimination due to any reason; Sowell, 2004).[7] Only if in an individual case no or mere bad quality information is given *and* the duty to acquire reliable individual level information was not possible to fulfill, it is rational and ethically acceptable, to use information from non-individual categorizations (Sesardic, 2005).

The presented results reflect past and current cognitive ability measures and they reflect a state of intelligence in historical and evolutionary development. Modernization may lead to narrowing gaps. Overcoming the most serious environmental obstacles as deficits in health care, nutrition and education could lead in Africa within one generation to a rise of 10 to 15 IQ. Globalization is likely to lead via cultural exchange and via mixing to a phenotypically and genetically more homogenous human species. No matter which factors are relevant for differences between human subgroups, environmental improvement is important, as effective support at different levels of abilities and ages helps to provide a more beneficial development and to enable a life in self-responsibility. This can be done by better nutrition, health care, more and better education for students and teachers, and cognitive training. The large IQ rise in the last decades in more progressing African countries as Kenya (Daley et al., 2003) supports positive expectations. More globally, interventions should

[7] Martin Luther King (1986/1963, p. 219): "I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character."

also successfully target broader human capital (achievement motivation, discipline, self-concept and conscientiousness). This will improve cognitive development but also helps to deal at all ability levels with given ability.

References

Baker, J. R. (1974). *Race.* Oxford: Oxford University Press.
Caldwell, J. C. (2002). The African AIDS epidemic. *Journal of Population Research, 19*, 173-190.
Cavalli-Sforza, L. (1997). Genes, peoples, and languages. *PNAS, 94*, 7719-7724.
Ceci, S. J., & Williams, W. M. (2009). Should scientists study race and IQ? Yes: The scientific truth must be pursued. *Nature*, *457*, 788-789.
Dagona, Z. K. (1994). Substance use and road traffic accidents among Nigerian commercial motor cyclists. *Ife PsychologIA, 2*, 81-93.
Daley, T. C., Whaley, S. E., Sigman, M. D., Espinosa, M. P., & Neumann, Ch. (2003). IQ on the rise: The Flynn effect in rural Kenyan children. *Psychological Science*, *14*, 215-219.
Dawkins, R. (2008/1992). *The extended phenotype.* Oxford: Oxford University Press.
Flynn, J. (2007). Arthur Jensen and John Stuart Mill. *Cato Unbound*, *November* (26). Retrieved September 12, 2008, www.cato-unbound.org/2007/11/23/james-r-flynn/arthur-jensen-and-john-stuart-mill.
Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (II, pp. 945-1017). Amsterdam: North-Holland.
Gonzales, P. (2000). *Mathematics and science in the eighth grade.* Washington: Center for Education Statistics.
Gottfredson, L. S. (2010). Lessons in academic freedom as lived experience. *Personality and Individual Differences*, *49*, 272-280.
Harrison, L. E. (2006). *The central liberal truth.* New York: Oxford University Press.
Hart, M. H. (2007). *Understanding human history.* Augusta: Washington Summit.

Hunt, E. (2011). *Human intelligence.* Cambridge: Cambridge University Press.
Kabou, A. (1991). *Et si l'Afrique refusait le developpement?* Paris: L'Harmattan.
King, M. L. (1986/1963). *The essential writings and speeches of Martin Luther King, Jr.* New York: Harper Collins.
Lindsay, R. (2010, September 13). Liberal race realism: Clearing up a few things. Retrieved June 25, 2012 from http://robertlindsay.wordpress.com/2010/09/13/liberal-race-realism-clearing-up-a-few-things.
Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality.* Athens: Washington Summit.
Murray, Ch. (2003). *Human accomplishment.* New York: Harper-Collins.
Nisbett, R. E. (2009). *Intelligence and how to get it.* New York: Norton.
Nyborg, H. (2003). The sociology of psychometric and bio-behavioral sciences. In H. Nyborg (Ed.), *The scientific study of general intelligence* (pp. 441-502). Oxford: Pergamon.
Oesterdiekhoff, G. W., & Rindermann, H. (2007). The spread of AIDS in developing countries: A psycho-cultural approach. *Journal of Social, Political and Economic Studies, 32*, 201-222.
Piaget, J. (1953/1936). *The origins of intelligence in children.* New York: Routledge.
Pinker, S. (2007). Groups of people may differ genetically in their average talents and temperaments. In J. Brockman (Ed.), *What is your dangerous idea? Today's leading thinkers on the unthinkable* (pp. 13-15). New York: Harper.
Rindermann, H. (2011). Increase of intelligence in childhood and youth. *Psychologie in Erziehung und Unterricht, 58*, 210-224.
Rindermann, H. (2012). Intellectual classes, technological progress and economic development: The rise of cognitive capitalism. *Personality and Individual Differences*, *53*, 108-113.
Rindermann, H., & Meisenberg, G. (2009). Relevance of education and intelligence at the national level for health: The case of HIV and AIDS. *Intelligence, 37*, 383-395.

Rindermann, H., Sailer, M., & Thompson, J. (2009). The impact of smart fractions. *Talent Development and Excellence, 1*, 3-25.
Rindermann, H., & te Nijenhuis, J. (2012). Intelligence in Bali – A case study on estimating mean IQ for a population using various corrections based on theory and empirical findings. *Intelligence, 40*, 395-400.
Rindermann, H., & Thompson, J. (2011). Cognitive capitalism: The effect of cognitive ability on wealth, as mediated through scientific achievement and economic freedom. *Psychological Science, 22*, 754-763.
Rindermann, H., Woodley, M. A., & Stratford, J. (2012). Haplogroups as evolutionary markers of cognitive ability. *Intelligence, 40*, 362-375.
Rushton, J. Ph. (2000/1995). *Race, evolution, and behavior.* Port Huron: Charles Darwin Research Institute.
Rushton, J. Ph. (2004). Placing intelligence into an evolutionary framework or how g fits into the r-K matrix of life history traits including longevity. *Intelligence, 32*, 321-328.
Rushton, J. Ph. & Jensen, A. R. (2005). Thirty years of research on Black-White differences in cognitive ability. *Psychology, Public Policy & the Law, 11*, 235-294.
Rushton, J. Ph., & Rushton, E. W. (2004). Progressive changes in brain size and musculo-skeletal traits in seven hominoid populations. *Human Evolution, 19*, 173-196.
Rushton, J. Ph., Skuy, M., & Fridjhon, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White engineering students in South Africa. *Intelligence, 31*, 123-137.
Segerstråle, U. (2000). *Defenders of the truth.* New York: Oxford University Press.
Sesardic, N. (2005). *Making sense of heritability.* Cambridge: Cambridge University Press.
Signer, D. (2004). *Die Ökonomie der Hexerei. [The economy of sorcery.]* Wuppertal: Peter Hammer.
Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjhon, P., & Rushton, J. Ph. (2002). Effects of mediated learning experience on

Raven's matrices scores of African and non-African university students in South Africa. *Intelligence*, *30*, 221-232.
Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist, 42*, 137-144.
Sowell, Th. (2004). *Affirmative action around the world.* New Haven: Yale University Press.
Wicherts, J. M., Dolan, C. V., & Maas, H. L. J. v. d. (2010). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence*, *38*, 1-20.

See also the *supplementary data file*.

Supplementary material for African cognitive ability: Research, results, divergences and recommendations: Method and further results

1. Study 1: Tanzania country sample

1.1 Method

The dataset used in this study was collected by Anke Müller[8] in Moshi, Tanzania, for her master-thesis at the end of 1999, beginning of 2000 (Müller, 2002).

1.1.1 Research participants

N=174 students of grades 8 and 9 attending a public school (public, but with fees, Majengo secondary school) in Moshi were tested with three cognitive ability tests. Moshi is a city in the North of Tanzania near to the Kilimanjaro (population 140,000 in 2000), with an economy primarily based on services, tourism, small industry, and agriculture. The city has various higher education institutions, including Mweka College, which offers courses in wildlife management, and Tumaini University, which offers courses in humanities, law, and social sciences. Due to Moshi's altitude (at around 800m) the average temperature is lower than in other Tanzanian cities. The city is said to be the cleanest of East Africa.[9] Compared to other regions (rural regions) the living conditions are relatively high, with personal incomes being above regional and country averages. The school, where testing occurred, was selected based on the willingness of the director to participate in the study. The environmental characteristics were generally above average, possibly biasing cognitive ability results in an upward direction.

59% of the sample were boys, and 41% were girls (for schools in Africa not atypical). Probably girls have to work more

---

[8] Anke Müller was 1999-2002 at University of Magdeburg student assistant in the psychology department of Heiner Rindermann. She lived for six month in Moshi for a practical training and for collecting data for her master-thesis.

[9] Retrieved January 20, 2012 from http://en.wikipedia.org/wiki/Moshi.

frequently at home. Mean age of students was 16.32 years (*SD*=1.05, range from 14 to 21 years). *N*=109 students were tested with the APM (tests explained below; age: *M*=16.10, *SD*=0.99, range from 14 to 21 years), *N*=88 students were tested with the MRT (age: *M*=16.55, *SD*=1.01, range from 14 to 19 years), and *N*=96 students were tested with the LPS (age: *M*=16.60, *SD*=1.02, range from 14 to 19 years).

1.1.2 Observations on school attendance, instruction and everyday life

Anke Müller (2002) gave also some remarks on casually observed school conditions: Classes were large (around 60 students per class and teacher). Teachers stressed rote learning, not comprehensive learning, and prioritized respect and obedience over independent thinking (see also Glewwe & Kremer, 2006). Examinations generally tested for correct reproduction, not understanding and transfer. Students had a good memory, paper notebooks were rarely used. Daily life was seen as slow and boring. According to her, in Moshi were no possibilities for entertainment and diversion.

1.1.3 Tests

Three different tests were applied:

1. *APM* (Raven Advanced Progressive Matrices; Raven, 1976/1943). The APM are a measure of fluid (knowledge-distant, culture-reduced) intelligence. They consist of two sets with 12+36 (48) figural tasks. In the first tasks a field with an empty corner has to be complemented out of eight given possible solutions below. In the later tasks a matrix with nine figures is given, one is missing; the missing one has to be found among eight possible solutions given below. The first tasks are very easy and nearly solvable by visual inspection. But for all following tasks reasoning abilities are necessary: Rules have to be detected (continuation, opposite/reflection, composition of single parts, becoming more, greater, smaller, 1:2 is as 2:4, addition, subtraction). At least some basic mathematics is indispensable (e.g., counting to 10 and

addition and subtraction), but the main focus is on the detection of rules and their application (i.e., reasoning).

The Raven Matrices tests are widely used in cross-cultural research, but, like other figural tests, they are not culture-free because: (a) amount of school education furthers intelligence including fluid; (b) different cultures differently weight school education, thinking, and cognitive stimulation; (c) different cultures differently further school education, thinking, and cognitive autonomy; (d) different cultures and societies differently further non-cultural environmental conditions relevant for cognitive development as quality of nutrition, health care, stimulation by modern life etc. But the Raven are specific-culture-free in the aspect that a specific language or a specific instruction in mathematics and other school subjects are not necessary to understand the test instruction and the tasks.

Compared to the SPM (Standard Progressive Matrices) the APM are more difficult. During the testing it became obvious that the APM were too difficult for the majority of students. According to Müller even some teachers have had problems to understand the items. Three "reasons" can explain the selection of the APM (and not of the more appropriate SPM): First, the supervisor of this study (Wolfgang Lehmann) and the author of this paper (HR, former colleague of Wolfgang Lehmann) were coming from an institutional tradition of high ability research. The SPM with about five to ten cognitively demanding tasks was not seen as a serious instrument to measure intelligence, at best usable in primary school for 6 to 10 year old students. For secondary school students only the APM were seen as conceivable. Second, international comparisons of intelligence and cognitive ability (Lynn & Vanhanen, 2002, 2006; Rushton, 2000, 1995) and studies in the tradition of Piaget (Bruner, Olver, & Greenfield, 1966; Luria, 1976/1974) were at the time unknown. Third, there was a widespread conviction, that people in developing countries are even *smarter* than people in developed ones, because the former have to deal with many difficult everyday problems (as organizing food, energy, water, repairing bicycles, machines etc.) which for the latter

are solved by others, experts or do not emerge due to better designed environments. Thus – so was the idea, stimulated by Rousseau – the people in developing countries would have an everyday life training in intelligence which furthers their cognitive development. The at-first-sight puzzling low results started the interest of the author (HR) in this research area.

2. *MRT* (Mental Rotations Test; Peters et al., 1995). The MRT uses 10-block, three-dimensional figures, which are correctly or wrongly rotated (that is a different figure). Testees have to find out which figures (out of four: two are correct) are after rotation identical or not with the first target figure. The MRT measures spatial ability, a subcomponent of general cognitive ability (due to time pressure as many other tests also somewhat processing speed). The norms for older students (older than 18, N=4) are somewhat too hard (in the German reference sample was for this age a mixture of Gymnasium and university students). Because only four of 96 students were older than 18, we did not correct the results for these students (the change would be negligible).

3. *LPS* (Leistungsprüfsystem; Horn, 1983). The LPS is a broad German intelligence test based on Thurstone's model of primary mental abilities measuring cognitive ability with 10 scales including various verbal tasks. For the study in Tanzania all scales requiring German language were left out. The following scales were used: LPS 3 (figural rows have to be continued, reasoning using figures, some mathematical abilities would be useful), LPS 4 (letter rows have to be continued, mathematical reasoning using the alphabet), LPS 7 (similar to MRT, but easier: rotated, but identical figures have to be find), LPS 8 (a geometric body is marked at one place, then the body is presented as an unwound net and the testee has to mark the right place), LPS 9 (all sides/facets of a geometric body have to be counted), LPS 10 (hidden figures in complex patterns have to be detected). The applied scales of the LPS measure fluid intelligence.

In summary, all tests measure culture-reduced, school-distant, fluid (as opposed to crystallized) intelligence, requiring only knowledge of letters and the alphabet.

1.1.4 Testing

The students were eager to be tested. Some wanted to be tested twice. Students were highly motivated. Students from other schools, who heard about the testing, came and wanted to be tested too, but regrettably, were rejected (unfortunately for training reasons, testing is a kind of training). Even teachers wanted to be tested. To prevent copying students used their arms and exercise books hiding their solutions from their neighbors. Every testee had his/her own desk and chair, but desks and chairs were somewhat noisy when moving (metal on stone). Moreover, classrooms had walls with open slots (where windows would normally be), outside noise could enter. All students and teachers spoke English, but test instruction was somewhat difficult. The tasks, also the APM tasks, were not perceived as having obvious objectives and the test instructor needed several examples and visual demonstrations (e.g. using buttons, geometrical bodies and sketches at the board) to explain the tasks. The pattern of solutions indicates that the testees understood the tasks (APM: no zero results, MRT: only 8% zero results, LPS: no zero results in all single scales). Due to time constraints not all tests were used for all students (APM: 65%, MRT: 56%, LPS: 56%).

1.1.5 Analysis and corrections

The *APM* used German norms from 1979 (Kratzmeier & Horn, 1980). Because the norms were about 20 years old in the year 2000, a Raven Matrices FLynn[10] correction of 2 IQ-points per decade was used (Lynn, 2009; Lynn & Vanhanen, 2002; Rindermann & te Nijenhuis, 2012). That means, a German sample, the benchmark here, would have solved in the year 2000 more tasks corresponding to an increase of 4 IQ points. Thus per decade 2 IQ-points, in sum 4 IQ points were subtracted.

The *MRT* norms used German age-appropriate means and standard deviations collected in 2001-2002 in the German state

[10] "FLynn effect", a combination of the names of the two men who rediscovered secular score gains in intelligence, namely Richard Lynn and James Flynn.

Sachsen-Anhalt at secondary schools (lower ability level) and Gymnasium (higher ability level) from Geiser, Lehmann and Eid (2008). Sachsen-Anhalt achieved slightly below average results in Germany (IQ 99 vs. German mean set at 100). Sachsen-Anhalt has a very small immigrant share among students. The selection of schools could be slightly biased towards higher levels of ability (usually, schools with higher ability levels tend to participate in test studies; the effect could be around 1 to 1.5 IQ points), but the Sachsen-Anhalt mean in past tests was below average – so no correction is necessary. The norms seem to be quite representative for the German ability spectrum. No FLynn-correction for the year 2000 was applied; it would only result in a minor upward correction of around 0.2 to 0.3 IQ points.

The *LPS* used German norms from 1960 (Horn, 1983). Because the norms were about 40 years old in 2000, an average FLynn correction of 3 IQ-points per decade was used (past FLynn-effects were larger then today ones in developed countries, mean of 4 to 2 IQ per decade; 12 IQ points subtracted; Lynn & Vanhanen, 2002).

Greenwich IQ correction: For international comparisons one common norm is necessary. Due to historical reasons (important cross-cultural tests were developed in Great Britain as the Raven Matrices, the first book on international IQ comparisons was published by a British researcher) the British norm serves as this standard. The German mean in international psychometric IQ and student assessment tests is IQ 99 (in Greenwich norm; IQ 99, Lynn & Vanhanen, 2006; IQ 99, Rindermann, 2007; IQ 99, Rindermann, Sailer & Thompson, 2009). Thus the results for Tanzania standardized in German norms have to be corrected with -1 IQ to come to the British norm used for international comparisons.

School enrollment correction: The results give only information for students – youth at school. Youth not at school will have a lower IQ:

1. Due to *regional selection effects*: In better developed regions more will attend school and in such regions will exist further for cognitive development favorable conditions, as better health

care and nutrition and more education among parents, higher wealth among parents usable to cover school attendance costs, better paid, better educated and more motivated teachers teaching at schools (and not doing other work).

2. Due to *family and individual selection effects*: Parents with higher education and IQ will send their children to better schools for longer periods of time, additionally wealth correlated with parental IQ is helpful or necessary for school attendance (fees, uniform, transport, books, meal, no use of the workforce of children; sometimes up to presents, supplementary, but necessary salary, or bribery for teachers).
3. Due to *school attendance effects*: Learning trains intelligence (e.g. Ceci, 1991, Rindermann & Ceci, 2009): If we want to generalize to total youth we need to include unschooled youth and their probably lower cognitive ability level. Usually per year attended at school the IQ rises about 3 IQ points (Cliffordson & Gustafsson, 2008: 3.7 IQ; Stelzl et al., 1995: 5.1 IQ; Winship & Korenman, 1997, overview study: 2.7 IQ; Hansen, Heckman & Mullen, 2004: 3.0 IQ; Falch & Massih, 2011: around 3.0 IQ). These values are all smaller than average IQ increases per one year in youth (around 5.62 IQ points; Rindermann, 2011) because the IQ growth could be traced back to school (institutionalized education) and non-school (from biological maturation to learning in everyday life). The older Winship and Korenman overview came to a rather unusual low gain. The best guess should be somewhat higher, around 3 IQ per year pure school grade effects. In Africa due to very large classes (up to 80 students), low teacher education, interruptions of school attendance, bad student health and insufficient learning equipment (e.g. missing boards; Glewwe & Kremer, 2006) the gain per school year is probably less, around 2.5 IQ per year.

In 2000 (Barro & Lee, 2000) 57.20% of 25 year old persons in Tanzania had (at least some) school education. Other sources report similar school enrollment data (e.g. US Department of Labor, 2009, p. 353; 57% school attendance, children 5-14, 2000).

Newer reports from Barro and Lee (2011) are more optimistic: The average percentage of 15- to 19-year-olds with no schooling is 21.7%, 14.7%, and 0.1% in years 2000, 2005, and 2010, respectively. But what means "no schooling"? Never been inside a school building? Or only for few months and years? No primary school qualification? It is difficult to decide which number stands for the population enrolled 2000 in secondary school eligible for testing. We averaged the three numbers for 2000 (57.20%, 57% and 78.3%; mean: 64.17% in school, 35.83 not in school). How many years the youth 15 to 19 year old not at school went to school in younger ages? Barro and Lee (2011) report for 2000 on average 4.61 years of schooling. Our students were tested in grades 8 and 9. Maybe the youth not at school went on average 3 years to school (the number has to be lower than the general average), thus they have missed compared to the tested sample on average 5.5 years of school. Assuming an increase of 2.5 IQ per school year we have to calculate the mean IQ for the total youth with this formula (for clarification more parentheses added than mathematically necessary): (measured IQ × .6417) + ((measured IQ-(5.5 × 2.5 IQ)) × .3583). Because in the older fractions of society the school attendance rate is even lower (we optimistically assume about 5% less), the total society IQ would be somewhat lower: Mean IQ for the total population can be estimated using this formula: (measured IQ × .5917) + ((measured IQ-(5.5 × 2.5 IQ)) × .4083).

Secular intelligence rise correction: The last decades' IQ rise (Tuddenham-Lynn-Flynn-effect, "FLynn-effect"; Tuddenham, 1948; Lynn, 1982; Flynn, 1984; Kaufman & Weiss, 2010) seems to be higher in developing African countries than in developed Western ones. So Daley et al. (2003) reported an IQ-rise (1984 to 1998) of 8.02 IQ points per decade for Kenya, Khaleefa et al. (2008) reported a rise of 2.90 IQ points per decade (1964 to 2006) for Sudan, Meisenberg et al. (2005) for Dominica (with around 87% Africans, GDP level is comparable to richer African countries as Congo, Namibia and South Africa, cross-sectional age comparison) of 5.10 IQ points, and Wicherts, Dolan, Carlson and Maas (2010a) in an overview on around six African countries for

adults 7 IQ per decade, but for children only around 1.4 IQ (both averaged 4.18 IQ). There is no reason why the FLynn effect for children should be smaller than for adults. Most probably the difference is a sample artifact. Older studies in South-Africa showed for Africans (compared to Europeans) no or only small increases during apartheid (Richter, Griesel & Wortley, 1989; te Nijenhuis, Murphy, & Eeden, 2011). The best guess would be the mean taken from the overview study of Wicherts et al. (2010a) with 4.18 IQ per decade. According to Lynn (2009), the IQ increase per decade between 1979 and 2008 was 2.14 IQ points in Britain. Thus the IQ increase in developing countries with African majority is 4.18 IQ/d per decade, or 2.04 IQ/d per decade larger than that in developing countries.

The students were tested end of 1999/beginning of 2000. To come to a result for the year 2010 we have to deal with the assumed larger IQ increase in Africa. Therefore we added 2.04 IQ.

1.2 Results

Table S.1 shows the raw values and standardized values for each test and age in Tanzania. The raw results (not standardized depending on age) do not correlate with age (with APM, MRT and LPS: r=.02, -.08, -.11) or grade (r=-.07, -.20, -.13). If controlled for grade or age the pattern remains similar. There seems to be *no maturation* and *no cognitive ability training* by school between age 15 and 18 years, at least not in fluid intelligence.

Table S.1. Means (in parentheses SDs and N) for the Tanzanian sample (raw values and original IQ depending on age).

	Age	Raw values	IQ in original norm
APM	15	15.50 (8.19, 28)	78.39 (11.28, 28)
(Raven, general figural,	16	13.52 (7.47, 50)	75.79 (9.35, 50)
fluid intelligence)	17	16.71 (7.96, 24)	79.86 (11.06, 24)
	18	13.57 (6.08, 7)	75.34 (6.55, 7)
MRT	15	5.54 (2.93, 13)	84.33 (9.72, 13)
(Mental rotation,	16	3.40 (2.66, 30)	77.14 (8.23, 30)
"space", fluid	17	3.97 (3.40, 38)	73.01 (11.99, 38)
intelligence)	18	4.13 (2.75, 15)	73.68 (9.12, 15)
LPS	15	16.83 (3.58, 13)	93.75 (7.14, 13)
(General, figural, fluid	16	16.22 (2.92, 30)	92.00 (5.66, 30)
intelligence)	17	16.63 (3.56, 38)	93.55 (6.97, 38)
	18	15.31 (4.58, 15)	91.08 (8.25, 15)
LPS3	15	16.77 (5.20, 13)	87.88 (12.45,13)
Figural rows	16	15.03 (4.97, 30)	83.00 (10.95,30)
	17	16.50 (5.86, 38)	88.16 (13.47,38)
	18	16.07 (5.31, 15)	87.50 (10.86,15)
LPS4	15	20.85 (3.60, 13)	95.38 (12.45, 13)
Letter rows	16	19.13 (5.22, 30)	90.25 (12.79, 30)
	17	20.16 (5.66, 38)	91.71 (12.18, 38)
	18	18.53 (6.02, 15)	88.50 (13.26, 15)
LPS7	15	10.08 (3.30, 13)	86.15 (10.08, 13)
Rotated figures	16	10.73 (4.83, 30)	87.00 (14.06, 30)
	17	10.87 (3.75, 38)	87.57 (10.79, 38)
	18	12.33 (5.84, 15)	91.50 (12.63, 15)
LPS8	15	21.46 (11.05, 13)	104.62 (13.18, 13)
Geometric bodies in nets	16	19.40 (8.85, 30)	104.50 (10.16, 30)
	17	16.84 (9.06, 38)	101.58 (11.05, 38)
	18	14.00 (9.54, 15)	97.00 (11.96, 15)
LPS9	15	13.23 (7.60, 13)	85.58 14.18, 13)
Count sides/facets of	16	14.73 (5.50 ,30)	87.00 (10.41, 30)
bodies	17	16.58 (5.56, 38)	91.32 (11.10, 38)
	18	12.27 (4.03, 15)	83.00 (5.99, 15)
LPS10	15	18.62 (4.31, 13)	102.88 (8.41, 13)
Hidden figures	16	18.30 (5.43, 30)	100.25 (11.73, 30)
	17	18.82 (4.96, 38)	100.99 (10.04, 38)
	18	18.67 (7.31, 15)	99.00 (13.26, 15)

Notes: APM: sum score of two subscales; *MRT*: sum score of two subscales. Age groups 14 (max. *N*=2) and 15 (max. *N*=33) combined. Age groups 18 (max. *N*=13), 19 (max. *N*=4) and 21 (max. *N*=1) combined. *Original IQ*: no correction for Germany due to deviation from Greenwich IQ, old norms (no German FLynn effect correction), no correction for school attendance rate (Tanzania) and FLynn effect between 2000 and 2010 in Tanzania. In the APM were solved at maximum 35 tasks (of 48 tasks, 73%). In the MRT were solved at maximum 14 tasks (of 24 tasks, 58%).

Table S.2 shows standardized results. IQ-all is the mean of the three intelligence measures. APM-IQ and MRT-IQ correlate with *r*=-.05 (raw: *r*=.12, *N*=36), APM-IQ and LPS-IQ *r*=.42 (raw: *r*=.46, *N*=36), MRT-IQ and LPS-IQ *r*=.43 (raw: *r*=.46, *N*=96). The correlations for normed test results are somewhat lower, especially because for the APM no age increase is assumed (one common norm).

Table S.2. Means (in parentheses SDs and N) for the Tanzanian sample.

	N	IQ of students in 2000 norm	IQ of youth 2000	IQ of adults 2000	IQ of youth 2010	IQ of adults 2010
APM	111	72.44	69.31	68.87	71.35	70.91
		(10.28)	(10.28)	(10.28)	(10.28)	(10.28)
MRT	96	74.94	70.37	69.73	72.41	71.77
		(10.73)	(10.73)	(10.73)	(10.73)	(10.73)
LPS	96	79.71	74.78	74.09	76.82	76.13
		(6.80)	(6.80)	(6.80)	(6.80)	(6.80)
IQ-all	171	75.04	71.11	70.56	73.15	72.60
		(9.05)	(8.90)	(8.89)	(8.90)	(8.89)

Notes: IQ in 2000 norm: corrected for Germany due to deviation from Greenwich IQ and old norms (German FLynn effect); *IQ of youth 2000*: further correction for (low) school attendance rate in Tanzania; *IQ of adults 2000*: further correction for lower school attendance rate of adults in Tanzania; *IQ of youth 2010*: further correction for stronger IQ increase in Tanzania; *IQ of adults 2010*: further correction for stronger IQ increase in Tanzania.

The mean cognitive ability levels of current youth and total society are seen at IQ 73. One caveat: Moshi seems to be compared

to entire Tanzania a better developed region with better supply of educational institutions. But a correction would be difficult – maybe around 2 to 5 IQ points. If there is any bias in the final IQ estimate, the deviation is upwards, not downwards.

2. Study 2: Cognitive ability of the blind and non-blind in South Africa and Austria

2.1. Method

The two datasets used in this study were collected by Patricia Studeny and Sonja Eherer in South Africa and Austria for their master-theses at the middle and end of 2008 (Studeny, 2008; Eherer, 2009). Patricia Studeny has made her practical training one year before in South Africa and has had contacts to a school for the blind and visually handicapped.[11]

2.1.1. Research participants

South African participants were N=153 students (N=45 blind, N=58 visually handicapped, N=52 normal-seeing) between the age 10 and 16 years (M=13.46 years, SD=2.07; 49.7% boys, 50.3% girls). Austrian participants (M=14.45 years, SD=2.21; 49.2% boys, 50.8% girls) were N=63 students (N=14 blind, N=25 visually handicapped, N=24 normal-seeing). 73% of the blind and visually handicapped were handicapped since birth or the first year of life (younger than 12 months), and 27% acquired their blindness or visual handicap later in life. The South African blind and handicapped persons attended special schools for the visually handicapped in Cape Town and Worcester (89%) or integrated classes (11%). The normal-seeing students came from public schools in different districts of Cape Town. The Austrian blind and handicapped persons attended special schools for the visually handicapped in Vienna and Graz. The normal-seeing students came from public schools in Vienna and Graz.

[11] The supervisor was Heiner Rindermann. External expert for research questions on intelligence of the blind was Dr. Jan te Nijenhuis, University of Amsterdam, The Netherlands.

The South African sample was diverse in terms of languages (Afrikaans 43%, Xhosa 40% and English 17%), evolutionary and ethnic background and ("race", "ethnicity", following the in South Africa usual categories: African 41%, Mixed/Coloured 34%, European 25%) and socioeconomic status (middle classes 38%, lower classes 40%, Township 22%; see Table S.3).

Table S.3. Sample characteristics of the South African-Austrian study (numbers).

Country	Evolutionary-ethnic background	Socioeconomic status	Eyesight			All
			Blind	Visual handicap	Normal	
South Africa	African	township	9	14	9	32
		lower class	5	12	6	23
		middle class	2	1	4	7
		all	16	27	19	62
	Coloured	township	1	1	0	2
		lower class	15	10	11	36
		middle class	2	7	5	14
		all	18	18	16	52
	European	lower class	0	1	1	2
		middle class	10	11	16	37
		all	10	12	17	39
		township	0	0	1	1
Austria	European	lower class	2	3	1	6
		middle class	3	7	12	22
		upper class	1	0	2	3
		all	6	10	16	32

Notes: 8 of the 63 Austrian students had immigration background (2 blind, 5 handicapped, 1 normal-seeing). From 31 Austrian students there was no information on socioeconomic status; "Coloured" means in South Africa "mixed" European and African ancestry (with minor influence of Indian and Asian ancestors).

Evolutionary-ethnic background and socioeconomic status are correlated (Spearman's r=.71, N=153).

2.1.2. Tests

The Wechsler Intelligence Scale for Children (WISC-IV) scales were used for which eyesight is not necessary (frequently used tests for cross-cultural comparisons based on figural material as Raven's Matrices or CFT were not applicable). All students were tested in a single testing session. The following constructs and scales (in South Africa in English, in Austria in German) were used:

The construct *working memory* was measured with the subscales "Digit Span" (testees have to repeat orally given sequences of numbers, either as heard or in reverse order) and "Letter-Number Sequencing" (testees are given in mixed order numbers and letters which have to be ordered in the mind and presented in this ordered way). The minimum required knowledge is to know numbers and numerical order, letters and alphabetic order. The scale is school-distant, similar to fluid intelligence. Both subscales correlate with r=.47 (α=.63).

The construct *verbal comprehension* was measured with the subscales "Similarities" (to find out what different words have in common), "Vocabulary" (to know words and their meaning), and "Comprehension" (to understand situations and concepts). This scale is school-near, based on knowledge (including language) acquired in family, everyday life and school. The three subscales correlate with $r_{S\text{-}V}$=.53, $r_{S\text{-}C}$=.62 and $r_{V\text{-}C}$=.72 (α=.83).

Both higher order scales represent fluid and crystallized intelligence. They correlate with r=.25 (α=.39 with two higher order scales, α=.73 with five subscales).

2.1.3. Analysis and corrections

The WISC-IV results were standardized using a German-Austrian-Swiss standardization sample (Petermann & Petermann, 2007). We used the published norms. The subscales were then averaged to the higher order scales *working memory* and *verbal comprehension*. The data for German WISC-IV standardization were collected in 2005 and 2006 in representative samples. We assume a small IQ increase (in German speaking countries) between 2005 and 2008 of

around 0.66 IQ points (equivalent to around 2 IQ points per decade). Therefore both, the Austrian as well as the South African IQs, were subtracted with 0.66 IQ points.

Greenwich IQ correction: The German, Austrian and Swiss means in international psychometric IQ and student assessment tests are IQ 99, 100 and 101 (mean IQ 100.00; Lynn & Vanhanen, 2006), adjusted IQ 99, 101 and 101 (mean IQ 100.33; Rindermann, 2007) and IQ 99, 100 and 100 (mean IQ 99.67; Rindermann et al., 2009). The grand mean is exactly 100.00. Thus the results expressed in German-Austrian-Swiss norms do not have to be corrected to be consistent with the British norms for international comparisons.

Selection bias correction: The results indicate that the South-African sample is twice positively selected: Firstly, all groups (whether categorized due to evolutionary-ethnic background or eyesight) show a much higher cognitive ability level than found in internationally comparing student assessment studies (in IQ-scale 63; Rindermann, Sailer & Thompson, 2009) or psychometric IQ test collections (IQ 72; Lynn & Vanhanen, 2006). As a very cautious correction we set the IQ mean of the sighted African-Europeans (people of European ancestry in South-Africa) on 100 (as their countries where their ancestors were coming from, corrected from 104.78, -4.78) and subtracted therefore from all groups -4.78 IQ. Secondly, within South Africa the blind seem to be positively selected or (very successfully) furthered. They were excluded from further estimations of the total South African mean cognitive ability level.

The *composition of the South African population* has to be considered: 79.5% have African roots, 9.0% European, 9.0% mixed (Coloured), 2.5% Indian/Asian (Statistics South Africa, 2011). We do not have reliable information on socioeconomic status comparable to our used categories. For total South African values we weighted our data according to evolutionary-ethnic background: ((African-IQ×.795)+(European-IQ×.09)+(Coloured-IQ×.09))/.975.

School enrollment correction: In 2010, according to Barro and Lee (2011), only 0.1% of 15-year-olds in South Africa have not

attended school (2005 identical). Average year of schooling is 8.5 years. Our sample is quite representative in school enrollment and needs no correction to generalize to total youth. A small percentage of students were 16-years-old, and not all South Africans of this age attend school; but a correction would be hypothetical and not large (about 1 to 2 IQ points). But in the older cohorts are shares without school education (e.g. between 40 and 44 years 11%). Barro and Lee report for the population above age 15 a no schooling rate of 9.3%. Our students were tested between age 10 and 16. On average they will have attended school for about 8 years (mean age 14 years). People not at school will have missed compared to the tested sample on average 8 years of school. Assuming an increase of 2.5 IQ per school year (see supplement paragraph 1.1.5) we have to calculate the mean IQ for the total population with this formula: (measured IQ × .907) + ((measured IQ-(8 × 2.5 IQ)) × .093).
Secular intelligence rise correction was done, and found to be small (differences in IQ rise between South Africa and Austria between middle/end of 2008 and 2010 are around 0.41 IQ).

2.2 Results

The raw values are presented in Table S.4. (pp. 140 – 141)

Table 8.4. Means (in parentheses SDs and N) for the South African and Austrian samples' raw values depending on age).

Scale	Age	Raw values			
		South Africa			Austria
		African	Coloured	European	European
Digit Span	10	13.55 (3.42, 11)	13.14 (4.22, 7)	12.67 (1.75, 6)	14.25 (2.75, 4)
	11	12.83 (1.47, 6)	13.71 (3.15, 7)	14.00 (2.12, 5)	13.00 (3.02, 8)
	12	14.25 (2.92, 8)	13.33 (2.50, 9)	14.00 (1.83, 7)	10.00 (3.65, 7)
	13	12.38 (2.92, 8)	14.38 (2.77, 8)	13.75 (2.63, 4)	14.25 (4.43, 4)
	14	11.30 (2.45, 10)	12.67 (2.16, 6)	14.40 (2.88, 5)	12.90 (1.91, 10)
	15	12.89 (3.18, 9)	13.17 (2.14, 6)	13.33 (2.80, 6)	12.33 (2.73, 6)
	16	12.80 (2.86, 10)	13.44 (2.19, 9)	12.83 (1.72, 6)	11.08 (3.27, 24)
Letter-Number Sequencing	10	10.00 (2.14, 11)	10.71 (1.80, 7)	12.83 (2.48, 6)	12.75 (2.75, 4)
	11	9.17 (2.86, 6)	11.43 (2.51, 7)	11.60 (1.52, 5)	11.88 (1.96, 8)
	12	11.50 (2.56, 8)	11.11 (2.47, 9)	14.43 (1.40, 7)	10.00 (3.92, 7)
	13	7.88 (3.23, 8)	11.50 (1.69, 8)	12.50 (0.58, 4)	11.25 (3.10, 4)
	14	8.00 (3.23, 10)	11.83 (2.48, 6)	11.20 (1.48, 5)	11.80 (1.81, 10)
	15	10.11 (3.30, 9)	10.83 (0.98, 6)	10.67 (1.75, 6)	12.50 (2.43, 6)
	16	9.70 (2.58, 10)	11.44 (1.42, 9)	12.00 (1.27, 6)	10.25 (2.59, 24)
Similarities	10	9.27 (1.49, 11)	8.43 (1.40, 7)	11.67 (1.51, 6)	10.50 (1.29, 4)
	11	9.33 (1.75, 6)	9.14 (1.57, 7)	11.00 (1.22, 5)	9.13 (1.96, 8)
	12	9.13 (0.64, 8)	9.56 (1.81, 9)	11.00 (1.83, 7)	7.86 (2.67, 7)
	13	7.75 (1.49, 8)	10.00 (2.39, 8)	12.00 (1.15, 4)	10.50 (3.11, 4)
	14	8.30 (1.77, 10)	10.00 (2.19, 6)	12.20 (1.79, 5)	9.90 (2.60, 10)
	15	8.00 (1.73, 9)	8.33 (1.21, 6)	12.50 (1.05, 6)	7.00 (2.68, 6)
	16	8.30 (2.16, 10)	10.33 (1.12, 9)	12.50 (1.05, 6)	8.58 (3.03, 24)

Scale	Age	Raw values			
		South Africa			Austria
		African	Coloured	European	European
Vocabulary	10	4.33 (0.79, 11)	4.43 (1.27, 7)	7.67 (1.37, 6)	10.00 (2.16, 4)
	11	4.83 (1.94, 6)	5.43 (2.44, 7)	7.20 (1.64, 5)	9.50 (2.35, 8)
	12	3.75 (1.39, 8)	5.00 (1.73, 9)	7.14 (1.21, 7)	6.43 (3.05, 7)
	13	2.88 (1.89, 8)	5.58 (2.13, 8)	6.50 (1.29, 4)	7.75 (2.50, 4)
	14	3.00 (1.89, 10)	5.33 (3.39, 6)	6.60 (1.14, 5)	8.90 (2.08, 10)
	15	3.11 (1.76, 9)	4.00 (2.37, 6)	6.50 (1.87, 6)	8.83 (2.14, 6)
	16	3.20 (2.25, 10)	4.67 (1.94, 9)	7.11 (1.27, 6)	7.92 (2.48, 24)
Comprehension	10	6.00 (2.41, 11)	5.86 (1.77, 7)	8.33 (1.37, 6)	10.75 (1.26, 4)
	11	6.83 (2.64, 6)	6.14 (3.18, 7)	8.40 (1.14, 5)	8.88 (1.81, 8)
	12	7.25 (1.39, 8)	7.33 (1.87, 9)	9.14 (2.04, 7)	7.71 (3.25, 7)
	13	4.75 (1.28, 8)	7.13 (2.53, 8)	9.50 (1.91, 4)	8.25 (1.50, 4)
	14	5.50 (2.37, 10)	6.67 (3.72, 6)	9.60 (2.07, 5)	9.60 (3.20, 10)
	15	4.89 (2.20, 9)	7.17 (1.17, 6)	10.00 (1.41, 6)	9.00 (3.10, 6)
	16	4.70 (3.79, 10)	7.33 (0.82, 9)	10.33 (1.21, 6)	8.17 (2.85, 24)

Table S.5. IQ-means in the South African-Austrian study.

Country	Evol-ethnic backgr.	Socioec. status	Eyesight			All
			Blind	Visual handicap	Normal	
South Africa	African (N=62)	Fluid (WM)	118.03 (8.90)	105.41 (14.40)	94.81 (8.12)	105.42 (14.28)
		Crystallized (VC)	73.59 (6.02)	72.34 (12.05)	84.02 (5.69)	76.24 (10.38)
		N IQ-Sum	16 95.81 (6.30)	27 88.88 (11.12)	19 89.42 (5.24)	62 90.83 (8.91)
	Coloured (N=52)	Fluid (WM)	121.17 (12.27)	110.23 (9.99)	105.65 (5.36)	112.61 (11.62)
		Crystallized (VC)	77.67 (10.53)	79.90 (9.37)	89.65 (4.35)	82.13 (9.92)
		N IQ-Sum	18 99.42 (10.11)	18 95.06 (8.21)	16 97.65 (4.55)	52 97.37 (8.13)
	European (N=39)	Fluid (WM)	121.74 (8.92)	119.09 (9.00)	108.28 (10.53)	115.06 (11.24)
		Crystallized (VC)	92.94 (5.23)	88.42 (3.32)	101.28 (3.82)	95.19 (6.94)
		N IQ-Sum	10 107.34 (5.58)	12 103.76 (4.50)	17 104.78 (4.89)	39 105.12 (5.02)
	All (N=153)	Fluid (WM)	120.16 (10.32)	109.81 (13.05)	102.55 (10.14)	110.32 (13.27)

Country	Evol-ethnic backgr.	Socioec. status	Eyesight			All
			Blind	Visual handicap	Normal	
South Africa	All (N=153)	Crystallized (VC)	79.66 (10.89)	78.11 (11.67)	91.40 (8.68)	83.07 (12.05)
		N IQ-Sum	44 99.91 (8.94)	57 93.96 (10.74)	52 96.97 (8.06)	153 96.70 (9.63)
Austria	European (N=63)	Fluid (WM)	114.84 (16.96)	102.50 (15.36)	110.13 (12.75)	108.15 (15.38)
		Crystallized (VC)	88.34 (7.75)	87.98 (13.06)	96.05 (13.73)	91.13 (12.79)
		N IQ-Sum	14 101.59 (10.81)	25 95.24 (12.84)	24 103.09 (11.65)	63 99.64 (12.33)
Both countries	All ancestries (N=216)	Fluid (WM)	118.87 (12.30)	107.58 (14.11)	104.95 (11.50)	109.69 (13.92)
		Crystallized (VC)	81.75 (10.83)	81.12 (12.87)	92.87 (10.67)	85.42 (12.78)
		N IQ-Sum	58 100.31 (9.35)	82 94.35 (11.35)	76 98.91 (9.70)	216 97.56 (10.54)

Notes: IQs according to norms from German-speaking countries with mean IQ 100 from 2005/2006 with small norm correction (0.66 IQ points subtracted for all). In parentheses the standard deviations are given. *WM*: working memory, fluid; *VC*: verbal comprehension, crystallized.

Blind and visually handicapped persons have their strength in working memory (here as indicator for fluid intelligence), not in verbal comprehension (mean difference favoring working memory: 31.79 IQ points vs. in the normal seeing: 12.08 IQ points). The more visually handicapped a person is, the stronger is his or her working memory (IQ 119, 108 and 105). This pattern is robust across both countries and all biologically and culturally distinguished subgroups. Controlled for country, subgroup and SES eyesight (1=blind, 2=handicapped, 3=normal) correlates negatively with working memory (r_p=-.43) and positively with verbal comprehension (r_p=.37), with total IQ only slightly negative (r_p=-.12). Blind and visually handicapped persons develop a stronger working memory as compensation for their visual handicap and reduced possibility to store information in written form. More serious visual impairment within the visually handicapped (rated for 39 persons) is correlated with a higher working memory (r=.40), but nearly not with verbal comprehension (r=.10). This is a strong support for the *malleability* of even a very basic cognitive ability!

The mean IQ of our sample is average in Austria (rounded IQ 100) and very high in South Africa (IQ 97). In this result is not considered the composition of the country, our sample is not representative and we have not included older persons. The pattern of results for the three subgroups in South Africa is as predicted by evolutionary, cultural or social theories: Europeans show the highest cognitive ability (White, IQ 105), then mixed persons (Coloured, IQ 97) and then Africans (Black, IQ 91). The "Black-White-difference" of 14 IQ points is nearly identical to the one reported for the US (15 IQ; Gonzales, 2000, p. 59, 61; Hunt, 2011, p. 411f.; Lynn, 2008, p. 88, 118).

The values of all groups are higher than expected and reported by other studies (see Tables S.7 and S.8), e.g. the European-Africans have a mean IQ of 105, 5 IQ higher than the average German, Austrian, Swiss, Dutch or British IQ! The sample seems to be positively selected. The Africans' IQ with 91 is also higher than the mean IQ of African-Americans (in Greenwich-norm about IQ 84). Within South Africa the blind have a higher

mean IQ than the non-blind. They seem to be positively selected or they benefit from special education programs and their quality.

In a regression analysis with general IQ (IQ-sum) as criterion and several theoretically important predictors, *evolutionary-ethnic background* is the strongest predictor (β=.42/B=4.80 IQ, which means that differences between African and Coloured or European persons are on average 4.80 IQ), then *socioeconomic status* (β=.24/B=3.00 IQ, e.g. from township to lower classes on average 3.00 IQ), then *test language* (corresponding or not corresponding to language spoken at home, β=.13/B=2.82 IQ), then *country* (β=-.13/B=-3.34 IQ favoring South Africa) and finally *eyesight* (β=-.11/B=-1.32 IQ favoring the blind). For working memory there is a very strong effect of *eyesight* (β=-.48/B=-8.00 IQ favoring blind). For verbal comprehension the strongest predictor is *evolutionary-ethnic background* (β=.38/B=5.60 IQ) followed by *eyesight* (β=.34/B=5.34 IQ favoring sighted persons). It is not clear whether the "test language" is a real effect of language-difference (as testees have problems understanding tasks and dealing with items due to different mother language), because testers could adept to any language problem in the single testing sessions. But the language-effect is in the crystallized scale stronger than in the fluid scale.

The frequently-given objection that "racial" differences (understood as biological, social or cultural ones) are only masked social differences is not supported because a) evolutionary-ethnic background (EEB) has an effect that is independent from SES, b) the effect of EEB is larger than that of SES, c) and EEB could affect SES via genes as "extended phenotype" (Dawkins, 2008/1992). Of course, asymmetric power structures and the history of apartheid in South Africa had created strong negative effects on environmental quality. But 18 years ago (and in other countries as the US much longer) these politics were stopped and even reversed. There is now "positive" discrimination (affirmative action; Sowell, 2004; Farron, 2010). More importantly, other peoples who were also discriminated against in the past (as the Jews in Europe and in the

US up to the middle of the 20th century, the Chinese/Japanese in the past in the US, today in Southeast-Asia; Sowell, 2004; Karabel, 2005; Farron, 2010) have not suffered from their maltreatment in the form of lower intelligence. But very often all these groups have suffered from suppression and even genocide.[12]

Table S.6. Corrected IQ-means in the South African-Austrian study.

Country	Original IQ	Selection bias correction	Composition correction	School enrollment and Flynn correction
South Africa	96.70	90.62	86.39	86.39 (y) 84.53 (a)
Austria	99.64	99.64	99.64	99.64 (y,a)

Notes: Selection bias correction: in South-Africa for Europeans taken the average IQ in their main corresponding countries (Netherlands and UK, for all in South Africa -4.78) and blinds not considered; *Composition correction*: shares of Africans, Coloured and Europeans considered, formula (Africans, Europeans, Coloured): ((84.32×.795)+(99.57×.09)+(91.50×.09))/.975; *School enrollment correction*: among adults share of unschooled population considered (among youth 2008 no unschooled population), formula: (86.39×.907)+((86.39-(8×2.5))×.093); y=youth, a=all incl. adults.

Considering selection bias, evolutionary-ethnic composition and school enrollment results in an IQ of 86 for the youth and of 85 for total society (see Table S.6). Both values are much higher than those found in former student assessment studies (in IQ-scale 63; Rindermann et al., 2009) or psychometric IQ test collections (IQ 72; Lynn & Vanhanen, 2006; IQ 77; Wicherts et al., 2010b). Most probably our rather small sample was positively selected and our corrections are insufficient. But our results also demonstrate a) the usual European-African difference and b) the malleability of a basic cognitive competence, working memory.

[12] Cognitive ability is only one aspect of life; of course, the life itself and the guarantee of human rights are more important.

The last gives hope for the large possibilities of cognitive training in countries at lower ability levels.

3. Study 3: Student assessment studies

3.1 Method

Rindermann, Sailer, and Thompson (2009) presented a combination and restandardization of PISA, TIMSS and PIRLS data. We have now newer data from PISA 2009. In the same way as in the 2009 paper we combined the data from the *Programme for International Student Assessment* (PISA 2000, 2003, 2006 and 2009, 15 years old students, reading literacy, mathematics, science, 2003 also problem solving), *Trends in International Mathematics and Science Study* (TIMSS 1995, 1999, 2003 and 2007, 4th and 8th graders, mathematics and science, Kazakhstan 2007 due to very differing data excluded), and *Progress in International Reading Literacy Study* (PIRLS 2001 and 2006, 4th graders, reading).

To form a common score, the results were averaged in the following sequence: (a) within one grade, year and study between different scales (e.g. within TIMSS 1995, 4th grade, across math and science); (b) within one year and study between different grades (e.g. within TIMSS 1995, across 4th and 8th grade); (c) within one study between different years (e.g. within TIMSS, across 1995, 1999, 2003 and 2007); (d) within different grade vs. age study approaches across TIMSS and PIRLS (TIMSS and PIRLS are studies done in grades, PISA is a study done in a single age group); and (e) between different study approaches (across grade and age approach studies: TIMSS-PIRLS-mean and PISA-mean). All averaging was done using *z*-transformations calculating means and standard deviations in countries which participated in all samples used for averaging (so *z*-formula are based on the same countries and over- or underestimation are avoided). Subsequently the *z*-results were re-normed using means and standard deviations obtained by simple arithmetical averaging of all three study results (SAS-scale, Student Assessment Scale, with M=500 and SD=100) as an orientation. The SAS values were finally transformed to the more usual IQ-scale, using Great Britain as reference country, SAS-

SD were simply transformed to an IQ-scale ("Greenwich-IQ", *M*=100, *SD*=15). Results are provided for N_c=93 countries (mean for all: *M*=90.25, *SD*=10.60, *Min*=62.82, *Max*=106.67). The results were not corrected for age (students older than in other countries), grade (students in higher grades than in other countries), sample quality (except for the exclusion of the outlier Kazakhstan in TIMSS 2007 and taking the well developed Chinese region Shanghai for total China by subtracting -5 IQ), school enrollment and cognitive ability rise.

Eric Hanushek and Ludger Wößmann took a similar approach, but selected other (older) student assessment studies and combined and standardized them differently. Hanushek and Woessmann (2009, p. 25f., A2ff., A13ff.) calculated the mean ability level of students in math and science using US NAEP-results and an OECD (Organisation for Economic Co-operation and Development) standardization sample for 77 countries with older and newer student assessment studies (1964-2003, FIMS, FISS, FIRS, SIMS, SISS, SIRS, TIMSS, PISA, PIRLS). The US NAEP-results (yearly tests in an intertemporally comparable way since 1969) were used to find a common comparison scale to combine data from different studies. We used their data and transformed them into the IQ metric (UK with IQ=100). Results are provided for N_c=77 countries (*M*=93.96, *SD*=8.58, *Min*=72.09, *Max*=107.53).

The third source are African student assessment data (1983-2007) for N_c=7 African countries rescaled by Richard Lynn and Gerhard Meisenberg (Lynn & Meisenberg, 2010a). They compared different standardization procedures and rejected regressions (too small standard deviations) and direct conversions (within-country standard deviation of Great Britain set at 15, underestimation of African cognitive abilities, probably due to school education problems) and decided to apply an adaptation to the international psychometric IQ standard deviations. This procedure results in about 2.32 higher mean IQ. They further discussed necessary corrections for selectivity among African countries (economically and educationally more advanced African countries participated),

lower school enrollment rates, and floor effects (in some studies the IQ acquired by guessing in multiple choice tasks would be 77 IQ points) resulting in a final correction of around 11 IQ (not applied here for Tables S.7 and S.8).

Finally, we report results from SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality, www.sacmeq.org). SACMEQ organizes student assessments in grade 6 in reading and mathematics in Southern and Eastern African countries (see Table S.7). SACMEQ I (1995 to 1998) measured reading literacy (N_c=7 countries), and SACMEQ II (1999 to 2004, N_c=14) and SACMEQ III (2007, N_c=15) reading and mathematics (Makuwa, 2010; Hungi et al. 2010). The single scales correlate within and across single surveys with r=.77-.99 (N_c=6 to 15). The sum of all three surveys has a reliability of α=.98. Zanzibar was merged here with Tanzania.[13]

Information on participation rates or school enrollment is missing. The study has not used an international norm. South African results are not reported separately for students with European, mixed or African ancestry, which could have been used for international standardization.[14] We tried to standardize the results by comparing them with the cognitive ability measures of Rindermann (see below), the only source with enough "data" ("data" – because estimates included). So the SACMEQ-results do

---

[13] Postscript 2013: There are two further local African studies which I did not include (when writing the paper simply not known): Monitoring Learning Achievement (MLA; literacy, numeracy and life skills in fourth grade 1999, N=11; Chinapah et al., 2000) and Programme d'Analyse des Systèmes Éducatifs (PASEC; French and mathematics in second and fifth grade, due to low comparability we took only mathematics, N=11; Conference, 2008). I will add them in further international studies. However, they do not use internationally anchored scales (only results for African countries). Therefore, they only help in within African comparisons.

[14] In an email from one of the researchers of the project, 31 March, 2011: "The 'new' South African government does not conduct analyses by race, color or ethnicity, but by provinces SES, rural/urban, or gender". However, Statistics South Africa (2011) reports populations statistics for "population group" ("African", "Coloured", "Indian/Asian", "White").

not reflect student assessment differences *between* developed and developing countries, *only* student assessment differences *within* African countries.

We combined the four data sets to one student assessment sum (SAS), restandardized and the standard deviations adapted to the study with the largest country sample (SACMEQ).

The SAS-sum values were corrected with the given data (if provided: school enrollment rate, students older than in international average or in higher grades than in international average). Data were not available for several countries and were estimated using the African mean. Data were not corrected for older persons having less education and FLynn effect. Studies on general ability changes in PISA, TIMSS and PIRLS are missing. Because in past studies school enrollment was lower than in more recently studies real competence increases could be masked. A correction for older persons would lead to somewhat lower results.

According to the psychometric intelligence paradigm student assessment scales, as far as they involve only knowledge retrieval and not novel problem solving, measure crystallized intelligence (e.g. Cattell, 1987/1971; Carroll, 1993). Their results reflect knowledge and depend on school education, but school education is not the only single determinant. Other determinants include family education, aspects of society, and genes (e.g. Wainwright et al., 2005).

3.2 Results

The results are presented in IQ metric (Greenwich IQ, M=100, SD=15) in Table S.7.

Table S.7. Student assessment results in international IQ norms.

Country	Rind.-collection/st.	Han.-W-collection/st.	Lynn-M-collection/st.	SACMEQ (Rind.-st.)	Comb.	Comb. corr.
Botswana	75.10	79.38	74.00	75.24	75.67	75.46
Ghana	62.82	79.79	78.20		73.92	63.13
Kenya				81.02	84.02	78.63
Lesotho				64.91	67.90	62.51
Malawi				60.59	63.59	58.20
(Mauritius)				(84.58)	(87.58)	(82.19)
Mozambique			82.20	71.93	77.07	68.61
Namibia				64.41	67.40	62.01
Algeria		88.06	82.77		78.06	74.02
(Seychelles)				(82.53)	(85.52)	(80.13)
South Africa	64.76	72.09	69.10	70.19	72.46	63.85*
Swaziland		91.72	87.10	77.18	79.69	74.30
Tanzania				76.95	79.94	74.55
Uganda				69.45	72.55	67.06
Zambia				62.05	65.04	59.65
Zimbabwe		87.36	77.80	71.93	76.60	76.36
Mean (SD, N)	67.56 (5.60, 3)	83.07 (7.25, 6)	78.74 (5.98, 7)	72.36 (7.63, 14)	75.43 (7.13, 16)	70.04 (7.80, 16)

Notes: Rind.-collection/st.= Rindermann collection-standardization; Han.-W-collection st. = Hanushek-Woessmann-collection-standardization; Lynn-M-collection st. = Lynn-Meisenberg-collection-standardization; Combined: not simply averaged, but restandardized. In the first four columns no corrections were applied. Corrected: for school enrollment rate, age, grade. *: value for South Africa probably an underestimation, see text.

The comparatively low ability level of South Africa is striking, e.g. compared to Nigeria, Tanzania and Mozambique, which are all much poorer countries (see also Table S.8). South Africa has participated in TIMSS 1995 8th grade, TIMSS 1999 8th grade, TIMSS 2003 8th grade and PIRLS 2006, and achieved average results in the student assessment metric (*SAS-M*=500, *SD*=100) *SAS*=340, 259, 254 and 302 (*SAS*-mean=289, transformed in IQ without any correction due to low school attendance and older than average student population and without UK-standardization IQ 68). Most probably, there is no down-biasing mistake in the South African TIMSS-PIRLS-samples. More probably there is (in international perspective) an overestimation. More plausible, the values of other African countries are (internationally) overestimated and in *within-Africa comparisons* the values of South-Africa are underestimated. For instance, because compared to other African countries more youth is attending school in South-Africa (lower selectivity). All countries compared, the corrections seem to be too severe for Ghana and South Africa. And student assessment studies could boost international cognitive ability differences compared to psychometric IQ test studies because effects of (lower) educational quality amplify competence differences in school-near (crystallized) tests compared to school-far (fluid) measures.

We recommend for further use a weaker correction (but a precise number is hardly possible). As provisional adaptation, we use the mean of corrected and uncorrected values as a final estimate for SAS. The validity of cognitive ability differences within Africa between countries is questionable. Further and more representative samples and up-to-date UK norms are needed.

4. Study 4: Prediction in regressions with the use of highly IQ-predictive variables

4.1 Method

In international comparison studies several variables have shown high correlations (initially independent of causal assumptions) with cognitive ability levels: general societal development, education,

wealth, and skin brightness. The United Nations Development Programme (UNDP) general societal development indicator (Human Development Index, HDI) includes also wealth and education, so the only two *different* measures *general societal development* and *skin brightness* were taken to predict cognitive ability.

4.1.1 Measures

We chose the two non-redundant variables with the highest correlations outside sub-Sahara Africa with cognitive ability (HDI and skin brightness). They stem from the two competing paradigms, nurture vs. nature. However, to have a statistical and predictive value any causal theory is not necessary or if one or different theories are chosen their veridicality is not relevant for the predictive purpose. The *Human Development Index* (HDI) is a highly general measure of human development used by the UN (consisting of life expectancy, years of schooling, and Gross National Income/GNI; UNDP, 2010). General societal development could positively influence cognitive development via nutrition, health, education and general stimulation (from media to social environment and culture). At the same time HDI reflects the impact of cognitive ability (e.g. Rindermann, 2008). Further factors behind them as culture and history or genes are supposed (see remark above: causal theory is not relevant). In our international data set (excluding sub-Saharan African countries and only measured data) HDI correlates with the sum of all cognitive measures (r=.75; N_c=107; combination of student assessment studies and Lynn and Vanhanen's updated IQ collection 2010). HDI correlates higher with cognitive ability than single education or wealth (r=.67, .53/.60, last logged wealth).

Skin brightness (or "skin reflectance", "skin color", any association between "brightness" as light intensity and mentally "brightness" are not intended[15]) was taken from Jablonski and

[15] We do not use the term "skin color", because color is not the relevant aspect, but high or low melanization. We do not use the term "skin reflectance" because

Chaplin (2000, pp. 74f.) and Templer and Arikawa (2006, pp. 124f.). Both indicators (high reflectance/brightness) correlate at $r=|.91|$ ($N_c=43$, mean: $\alpha=.95$). Skin brightness could not have a causal effect on cognitive ability, though some researchers assume pleiotropic effects of genes (Jensen, 2006; Ducrest, Keller & Roulin, 2008). Skin brightness is an indicator of evolutionary history and is highly predictive at the international data level (excluding sub-Saharan African countries and only measured data: $r=.75$, $N_c=82$). Within societies in the 16th to 20th century, skin brightness was associated with different treatment of people (slavery, apartheid and other physical and psychological maltreatment of African people). Skin brightness is internationally highly correlated with history (e.g. colonialism). A biological explanation could be conceivable only as a net variable corrected for the most important remaining effect of history on intelligence (HDI) (see remark above: causal theory is for prediction not relevant). The two predictors correlate among each other ($r=.70$, $N_c=78$).

Cognitive ability (the criterion) was measured as a sum value based on *student assessment studies* (IEA-Reading, IAEP-II, TIMSS, PISA, PIRLS, SACMEQ, 1991-2009, partially corrected, if data given, for age, grade and school enrollment rates) and *psychometric intelligence test studies* (updated: Lynn & Vanhanen, 2006; Lynn & Meisenberg, 2010b; Lynn, 2010; only measured data, combination by Heiner Rindermann). The Hanushek and Woessmann (2009) and Lynn and Meisenberg (2010a) collections were not considered here (this results for Swaziland in a larger divergence between SAS and cognitive ability sum).

Within African countries (defined as all countries with African majority populations, $N_c=52$) the results were compared to the results of *study 1 to 3* (including student assessment results and SACMEQ), *Lynn's updated data* (combination by Heiner

---

"reflectance" is not the correct evolutionary association (skin didn't become "reflecting" as white to protect against sun but lost melanization to enable more vitamin D synthesis in regions with less sunlight).

Rindermann, wrong data for Equatorial Guinea deleted and now as estimated; see Hunt, 2011), and a list of *intelligence levels for African countries* provided by *Wicherts, Dolan* and *Maas* (2010b; Table 5, p. 14, penultimate column). Richard Lynn has estimated data for countries without IQ data by taking the averages of neighbor countries with similar populations. We assumed that countries without data in IQ and SAS generally have non-beneficial environmental conditions (e.g. civil war in Afghanistan and Somalia) and subtracted 4 IQ points (similar Rindermann, 2007). Lynn presents his data in the Greenwich 1979 norm. If we assume for Africa a stronger IQ-rise in the past decades (see paragraph 1.1.3) we need to add per decade 2.04 IQ resulting in a plus of 6.32 IQ points for the year 2010.

4.1.2 Comments to two datasets and on plausibility

Wicherts' et al. data (2010b)

Wicherts et al. standardized Raven data for the year 1979, but it is less clear what was done for the other tests. Taking Wicherts et al. data as data normed for 1979 will lead to serious theoretical problems. If we take their own 82 IQ estimate (in my reanalysis 78) and correct it using their (highly plausible) data for the FLynn effect (2.04 IQ/decade), then the estimated value for 2010 would be IQ 88 (starting 1979 with IQ 82, or 84 starting 1979 with IQ 78). This would mean that Africans' IQ in Africa would be 4 IQ points higher than that of African-American in the US (or the same). All intelligence theories assume a positive impact of nutrition, health, education and global environmental stimulation on cognitive development. No one doubts that these conditions are on average better for Africans in the US than in sub-Saharan-Africa. And many (about half the researchers; Snyderman & Rothman, 1987) assume that genes play a role. Following Cavalli-Sforza (1997, p. 7724), African-Americans have on average 30% of their gene pool from Europeans. If (yet unknown) genetic factors are relevant for the African-European intelligence-difference, then this admixture will also lead to a higher cognitive ability level of Africans in the US than in sub-Saharan-Africa. Finally people in the US will have

more test-experience. If any of these theories are correct (no scientist is known who denies them all), the African-American IQ should be substantially higher than the sub-Saharan-African IQ.

We have many reliable sources for African-American IQ estimates. Therefore, it is logically inescapable that something in the Wicherts et al. data is divergent (in data selection, standardization, or corrections) or the FLynn-correction is wrong (but we have convincing data from different sources and well-founded theories to explain that the IQ is rising in Africa). This would also be the case when a somewhat higher African-American value (e.g. IQ 87 or 88) would be assumed. Finally, how Wicherts et al. (2010b) arrived at an IQ of 82 for Africans in sub-Saharan Africa is not clear.

My own averaging of Wicherts et al.'s (2010b) data (Table 5) yields an IQ of 78.[16] For student assessments (their Table 1) Wicherts et al. report (frequently written TIMSS as "TIMMS") a mean IQ of 82. Depending on the standard deviations used, this can be a statistically correct result for students (youth at school), but most probably this is an overestimation for youth and adults because school enrollment was not considered. At the end they write (p. 16): "Our systematic review of the average performance of Africans on the Raven's tests suggested that the average IQ of Africans in terms of western norms on the Raven's is approximately 78 (UK norms) or 80 (US norms). Thus, it can be concluded that the average IQ test performance of Africans in terms of UK norms is around 81." How they got from 78 to 81 is not explained. What mathematical procedure stands behind "thus"? Even simple (and incorrect) averaging across different norms and using a further numerical result one line above (81.8) will end in 79.93, not in 81.

Brouwers, Vijver and Hemert's data (2009)

One further existing psychometric IQ collection study (with Raven's Progressive Matrices) is from Brouwers, Vijver and

[16] I also gave the Wicherts et al. paper to a university student of psychology to recalculate the mean ability level in Africa. He came to the same result (IQ 78).

Hemert (2009). They also provided data for Africa. This study was not included because the final results table showed many highly implausible results such as IQ 126 for Norway, 57 for Syria, 115 for South Africa (compared to 67 for Nigeria), for East Germany 81 (compared to 110 for West Germany), 117 for Romania (compared to 96 for Slovakia). The deviations from other cognitive ability sources are not comprehensible and we do not have any assumption or formula with which we can correct them.

Symen Brouwers (e-mail from 26 January 2012) attributes the problem to convenience sampling in original data sets, on which the meta-analysis is based on, which was not always possible to control for (e.g. no data given for SES and ethnicity).

Plausibility

As the reader can see here and in all further corrections, international intelligence research is guided by arguments based on plausibility, which is based on rational integration of information and theories from different sources. All the numbers presented here are not pure numbers "found" somewhere in the environment, but are estimates of the true scores of cognitive ability. They present the final result of a long chain of decisions (inclusion criteria, standardizations, corrections). This is true for Hanushek, Lynn, Rindermann and Wicherts data.

4.1.3 Country selection

All sub-Saharan African countries on the continent were considered. The northern countries (from Mauritania to Eritrea and Ethiopia) include peoples with some Indo-European genetic heritage (e.g. Tuareg) and smaller populations with stronger Indo-European background; the same is true for Zanzibar (Arabian influence). Following the settlement frontier, in 2011 Sudan split in a northern (Arabic, Muslim) and southern (African, Christian-animist) region. Our Sudanese data are estimated prior to the split. Madagascar's population consists of African and Austronesian ancestry. There is no hint that both groups differ in education and cognitive ability, so Madagascar was included. The following Caribbean countries with African majorities were included:

Bahamas (85% African ancestry), Barbados (93%), Dominica (87%), Haiti (95%), Jamaica (91%), and Saint Lucia (83%) (percentages according to CIA, 2012).

4.1.4 Analysis

The formulas to predict cognitive ability levels for African countries were developed in country samples *excluding* sub-Saharan African countries and using only *measured* data for cognitive ability. Measured and (by HDI and skin brightness) predicted mean cognitive ability values highly correlate in the outside Africa sample (r=.79, N_c=111), higher than the two singles predictors with national cognitive ability (both, HDI as well as skin brightness, with measured IQ r=.75). The means are quite similar in the same country sample (measured: 90.94 IQ, predicted: 91.30, N_c=111 countries), but the standard deviations of the predicted values were too small (e.g. for the sum measured values' SD=7.83, predicted values' SD=5.24). This would result in an underestimation of (comparatively) high ability nations (e.g. Finland, 102 vs. 98) and an overestimation of (comparatively) low ability nations (e.g. Libya, 83 vs. 89). Thus the standard deviations (enlarged to the empirical ones, HDI: 7.83/5.84; and SkinBright: 7.56/5.66) and the means were slightly corrected resulting in identical means and standard deviations and a non-changed correlation (r=.79, N_c=111) and smaller differences at the tails (e.g. now Finland IQ 102 vs. 100, now Libya IQ 83 vs. 88).

4.2 Results

HDI and skin brightness similarly predict cognitive ability (in countries excluding sub-Sahara Africa: $r_{HDI\text{-}IQ}$=.75 [.746], N_c=107, $r_{SB\text{-}IQ}$=.75 [.748], N_c=82). In a common regression analysis, their predictive validity is also similar ($\beta_{HDI \rightarrow IQ}$=.47, $\beta_{SB \rightarrow IQ}$=.44, N_c=78). Results of the estimations as well as measured results of different sources are presented in Table S.8.

Table S.8. IQ values for countries with African majority.

Country	S 1-2	SAS-k	Lynn 79	Wich	Rind	HDI-IQ	SB-IQ	HS-IQ	Diff
Angola			(68)		(64)	69	68	68	4
Bahamas			(84)		(80)	94		94	14
Barbados			80		80	94		94	14
Benin			(70)		(66)	71	68	69	3
Botswana		76	71		76	84	73	78	3
Burkina Faso			(68)		(64)	62	65	64	0
Burundi			(69)		(65)	61	68	64	-1
Cameroon			64		64	72	64	68	4
Cape Verde			(76)		(72)	77		77	5
Cent. Afr. Rep.			64	71	64	63	66	64	1
Chad			(68)		(64)	61	65	63	-1
Comoros			(77)		(73)	70		70	-3
Congo (Brazz.)			64	78	64	74	70	72	8
Cote d'Ivoire			71		71	68	71	70	-1
Djibouti			(68)		(64)	69	73	71	7
Dominica			67		67				
Eq. Guinea			(64)		(60)	78	73	75	16
Eritrea			76	85	76		71	71	-4
Ethiopia			69	69	69	64	70	67	-2
Gabon			(64)		(60)	85	68	76	17
Gambia			64		64	68	62	65	1
Ghana		69	71	73	65	73	68	70	6
Guinea-Bissau			(67)		(63)	61	66	64	1
Guinea			67		67	64	64	64	-3
Haiti			(67)		(63)	69		69	6
Jamaica			71		71	88		88	17
Kenya		81	72	80	75	73	69	71	-4
Lesotho		65	(67)		64	70	68	69	5
Liberia			(67)		(63)	62	67	65	2
Madagascar			82	82	82	71		71	-11
Malawi		61	60		60	67	66	67	7
Mali			68	74	68	62	72	67	-1
Mauritania			(76)		(72)	71	79	75	3

Country	S 1-2	SAS-k	Lynn 79	Wich	Rind	HDI-IQ	SB-IQ	HS-IQ	Diff
Moz'bique		73	64		74	61	63	62	-12
Namibia		65	72	74	69	82	67	75	5
Niger			(69)		(65)	59	68	64	-1
Nigeria		76	69	84	76	70	67	68	-8
Rwanda			76		76	67	68	68	-8
São Tomé & Pr.			(67)		(63)	74		74	11
Senegal			71	66	71	69	64	67	-4
Sierra Leone			64	91	64	63	68	65	1
Somalia			(68)		(64)		68	68	4
South Africa	85	68	72	77	68	82	75	78	10
Saint Lucia			62		62				
Sudan			77		77	67	71	69	-8
Swaziland		77	(68)		68	75	70	72	4
Tanzania	73	77	72		74	68	66	67	-6
Togo			(70)		(66)	70	68	69	3
Uganda		70	73	84	72	70	64	67	-5
Zaire (C., Kins.)			65	76	65	58	68	63	-2
Zambia		62	71	79	68	68	64	66	-2
Zimbabwe		76	72	82	75	51	68	59	-16

Notes: Countries with African majority populations. *S1-2*: Studies 1 and 2 using psychometric tests, values of youth and adults averaged (reference year 2010), both samples probably positively selected; *SAS-k*: student assessment studies from 1964-2009 combined, standardized and corrected and non-corrected averaged (sources Rindermann, Hanushek & Woessmann, Lynn & Meisenberg and SACMEQ); *Lynn79*: measured (and in parentheses estimated) psychometric intelligence test data by Lynn & Vanhanen (updated 2010; reference year 1979); *Wich*: measured psychometric intelligence test data by Wicherts et al. (2010), Raven reference year 1979, the others probably 1979 and younger, UK norm; *Rind*: measured (and in parentheses estimated) psychometric intelligence (reference year 1979) and student assessment test data by Rindermann; *HDI-IQ*: IQ predicted by the use of the HDI (SD corrected); *SB-IQ*: IQ predicted by the use of skin brightness (SD corrected); *HS-IQ*: IQ predicted by the use of HDI and skin brightness (mean); *Diff*: difference between measured (or estimated) cognitive ability (Rind) and by HDI and skin brightness together predicted IQ.

The student assessment sum correlates with Lynn-IQ (r=.35, N_c=14) and Wicherts-IQ (r=.55, N_c=8) and Rindermann-IQ (r=.79, N_c=14). In the same country sample of eight countries the correlation with Rindermann's data is the highest (r=-.10, r=.55 and r=.80, N_c=8). Due to the small sample size this result should not be taken as reliable.

Our measured and estimated IQ ("Rind"; measured: test data given; estimated: no test data given, estimated by neighboring countries' test results) for African countries (countries with African population majority) correlates with the predicted IQ values with r=.38 (N_c=50). Contrary to expectations, the correlations are for countries with only estimated values higher than the measured values (r=.62, N_c=19, vs. r=.31, N_c=31). The means of measured and estimated (m/e) IQ values and of predicted (pr) IQ values are very similar ($M_{m/e}$=68.51, $SD_{m/e}$=5.73, M_{pr}=70.08, SD_{pr}=7.12).

Nevertheless, there are remarkable divergences regarding single countries: There are large overpredictions for countries, where we have only IQ estimations, but no empirical IQ data (Bahamas, Equatorial-Guinea, Gabon). Probably empirical test data will show higher results than previously estimated (by neighbor countries and subtracting 4 IQ points for assuming non-beneficial conditions in countries with no IQ estimates). Other countries have according to their general development level (HDI) better chances for a high IQ than they actually seem to achieve (Barbados, Jamaica, South Africa). Maybe these three countries profited in society's development from the endurance of past British and Dutch institutions, mineral resources, and international investments. Others such as Zimbabwe and Mozambique are less developed than expected by their measured cognitive ability level, perhaps because they suffer from a past history of civil war and unfavorable government.

Within the sample of African countries, the single predictor skin brightness is compared to the single predictor HDI: Which is the better predictor (less deviation from the measured and estimated values)? For the predictions with corrected standard deviations the mean squared difference from original IQ for the HDI-predictions

is D^2=76.14 (measured: 72.08, estimated: 84.57), but in the same country sample (N_c=40) the analogous difference for the skin brightness-predictions is D^2=31.09 (measured: 29.38, estimated: 34.63). This means that a) measured IQ data seem to be more accurate than estimated and that b) skin brightness is a better predictor for IQ within African countries than general human development (wealth, health and education). This is independent from method (standard deviations of predictions not corrected or corrected; measured or estimated IQ data). It should be noted that the formulas were obtained in countries outside of sub-Sahara Africa.

The IQ and student assessment data from Lynn and Wicherts were compared: In the same sample of N_c=17 countries, the Lynn-data correlate more highly with the predicted IQ than the Wicherts data (r=.34 vs. r=-.02). The mean squared deviance from predicted IQ is smaller for the Lynn-IQ (D^2=29.86 vs. 149.50). If we adjust the Lynn-1979-norm to 2010 (FLynn effect, +6.32 IQ points), Lynn's 2010 and Wicherts's (reference year not for all tests clear) means are very similar (same sample, N_c=17, Lynn 1979, mean IQ 70, Lynn 2010, mean IQ 77, Wicherts mean IQ 79). SACMEQ data correlate with the predicted IQ with r=.16 (N_c=12).

4.3 Theory added: HDI and skin brightness

HDI measures general social development (content: wealth, education and health). Its statistical effect stands for broad environmental effects on cognitive ability and of cognitive ability on environmental quality. Skin brightness measures the degree of melanization as an evolutionary adaptive response to sunlight intensity and its difference between the tropics and the polar regions (balancing protection against sunburn and skin cancer versus Vitamin D production). Skin brightness itself cannot cause or depend on cognitive ability, but is theoretically an indicator and empirically, after correcting for HDI, a marker of evolutionary history. It has of course also a historically associated political meaning (related to past practices of colonialism and racism), but their possible aftereffects are covered (at least predominantly) by

the HDI indicator. However, it would not be correct to assume that colonialism has had only or even mainly negative effects on education and cognitive ability in colonized countries. The colonial era also brought schools, institutions and values supporting education (Acemoglu, Johnson & Robinson, 2001; Harrison, 2006). Their acceptance and effect was not independent from local receptivity (e.g. for Nigeria: Sowell, 2004).

In a *radical environmental interpretation* both variables (HDI and skin brightness) and their statistical effects would stand for environmental effects. Skin brightness effects would be a social effect connected to a superficial attribute, e.g. via racism, colonialism and similar negatively valued processes. More tamely, they would represent differences in geographically-given chances to develop beneficial environmental conditions (e.g. for Haiti; Diamond, 2011).[17]

In a *radical genetic interpretation* both variables and their statistical effects stand for genetic effects. HDI would be an expression of hidden genes and their work in society and history.[18] If we *try* to come to an answer from an epistemic view, using data and theories and their scientific integration, we come to a more (as we hope) balanced and veridical view: If HDI, skin brightness, and cognitive ability are correctly measured, skin brightness is the slightly better empirical predictor. This does not exclude environmental effects (smaller effects are not zero effects). Evolutionary theories (e.g. Rushton, 2000, 2004) are only very roughly operationalized by skin brightness. It would be better to have genetic markers such as haplogroups (Rindermann, Woodley & Stratford, 2012) and, of course, if they were known, intelligence coding genes and their frequencies across different peoples (Deary,

---

[17] Some scientists or persons, who feel that they themselves represent something like science or intellectual engagement, believe to sail with the tailwind and frequently use in "scientific" disputes also political or other researcher demonizing "arguments" (see Segerstråle, 2000).

[18] Many see such positions as politically unacceptable: Such a position, explicitly expressed for differences between evolutionarily or culturally defined human subgroups, would resemble old racist policies.

2012). Only by assuming that skin brightness effects are fully covered by environmental factors, which actually have an effect on intelligence and which are not measured by general social development (HDI), an only-environment-theory could be true. There is not much evidence that this will be true (see above remarks on possible effects of "extended phenotype", "racism", "colonialism" and "suppression"). The coherent empirical pattern (stability of intelligence differences in various measures across time, societies and countries) can be more parsimoniously explained by the acceptance of at least partly genetic effects. However, only when the intelligence coding genes are found can this be shown beyond reasonable doubt. For furtherance approaches genetic theories are less relevant, because even genetically determined phenomena as myopia can be corrected by environmental interventions.

5. Additional test like evidence

Drawing a person is a kind of cognitive ability test and was widely used in developing countries for measuring intelligence (Goodenough Draw-A-Man test; Goodenough, 1926; Richter et al., 1989; Colom, Flores-Mendoza & Abad, 2007). In the Piagetian tradition the conceptional quality of drawings (completeness, correctness of number, proportion, and perspective, not psychomotor or artistic quality) depends on developed cognitive stage. According to Keller (2010) Nso-grandmothers (Cameroon) draw people as cephalopods (head, arms, legs, feet, no body, no hair, no face, no eyes, no nose, no mouth). Nso-mothers add bodies and hair and some face details. Heidi Keller argued against the use of drawings as an indicator of intelligence because, in her view, they would represent only neutral cultural differences which wait for scrutinized research and interpretations (however, not given). Humans as cephalopod-drawings are found in Western samples regularly from age 2 to 5, rarely later (Stern, 1952/1914, p. 319), which would represent pre-operational thinking in Piagetian theory (age 2 to 6 years). Translated in adults' IQ, this represents an IQ of around 50 for Nso-grandmothers, and a somewhat higher IQ for

Nso-mothers (around 60-70, Piagetian theory is here less precise; however, a further hint for a FLynn-effect).

Piagetian tests were applied in cross-cultural research (Bruner, Olver & Greenfield, 1966) and the results indicate delayed cognitive development. Furthermore, formal operations, the highest stage, was rarely reached, and concrete operations, the second highest stage, was not always attained. In Western samples, the formal operational stage is usually acquired at age 11/12, and the concrete operational stage at age 7. Translated in adults' IQ, this represents an IQ around 70 or 50. However, systematic cross-country comparisons, newer samples, and standardizations to newer British norms are missing.

6. Final remarks

All studies on cognitive ability in Africa suffer from problems of low to unknown sample representativity. Corrections as assumed FLynn and school absence effects are always somewhat speculative. Given the quality of the data it is not possible to come to a truly precise result. There are further arguments indicating overestimation (student assessment studies not corrected for older ages) or underestimation (less test-wiseness). International standardizations are based on with researchers varying formulas and assumptions. Researchers themselves discuss different true values ranging between 68 and 78 (Lynn & Meisenberg, 2010a) and Wicherts et al. (2010b) show mean results between 77 and 81 or 82 (82 is their own final best guess). Not only the corrections are based on plausibility arguments, but also the selection of studies. Even phenomena such as "school attendance" become fuzzy on further consideration: Is in "school attendance" an attendance included, when young boys and girls had been at school only 50% of the days with school instruction at school? Many corrections are tentative, making the research even for researchers themselves confusing and results intransparent. So the discrepancy between Wicherts et al. and Lynn results could be at least partially influenced by a disregarding of 30 years different IQ rise in Europe and Africa.

Certainly it is influenced by considering or not considering school attendance rates.

If we do not count with larger differences in test-wiseness effects (Europeans would have more test-wiseness and this boosts their measured IQ, but not their real cognitive ability), the present 2010 data tend more to overestimate than underestimate mean IQs (due to positively selected samples and optimistic FLynn-assumptions).

Future research should examine representative samples of the ages 10 to 70 (the coming and present workforce including political leaders); samples representative of the full range of school education, including the share without or with only little education; the use of fluid (school-distant, culture-reduced) and crystallized (school-near, student achievement) cognitive ability tests[19]; up-to-date norms from Great Britain; and, as shown by Rushton, Skuy and Fridjhon (2003) and Skuy et al. (2002), as test repetition and training improve test results and African samples may have less experience in tests and more generally in cognitive tasks, surveys with a short test and task training one or two weeks before for all tested persons.

Five necessary clarifications: First, even though the better prediction by skin brightness raises the possibility of genetic effects[20] any genetic effect does not rule out the possibilities of furtherance of cognitive abilities by environmental means (from improved nutrition and health care to cognitive training, see our working memory plasticity results). Second, to substantiate genetic effects beyond doubt we need to find the genes for intelligence. Third, environmental effects do not exclude genetic effects beneath and behind them. Fourth, all results are based on data which could be improved in scope and quality. Fifth, the individual counts. There are significant overlaps between human subgroups. We are

[19] The differences in crystallized intelligence and knowledge should be larger than in fluid.

[20] For many this will be a result which makes them feel uncomfortable and which stimulates some of them to attack researchers in an unscientific and unethical way.

strictly opposed to any political misuse of our research. There are smart Africans and dull Europeans. From an ethical viewpoint, we have to treat persons as individuals and we have to use the possibilities of intervention at all ability levels for all people.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review, 91*, 1369-1401.

Barro, R. J., & Lee, J.-W. (2000). *Barro-Lee data set. International data on educational attainment: updates and implications.* Boston: Harvard University. Retrieved November 18, 2004 from www2.cid.harvard.edu/ciddata/barrolee/readme.htm.

Barro, R. J., & Lee, J.-W. (2011). *Barro-Lee educational attainment dataset.* Cambridge: Harvard, Version 1.2, 09/11. Retrieved January 12, 2012 from www.worldbank.org.research/growth/ddbarle2.htm.

Brouwers, S. A., Vijver, F. J. R. v. de, & Hemert, D. A. v. (2009). Variation in Raven's Progressive Matrices scores across time and place. *Learning and Individual Differences, 19*, 330-338.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. (1966). *Studies in cognitive growth.* New York: Wiley.

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies.* New York: Cambridge University Press.

Cattell, R. B. (1987/1971). *Intelligence: Its structure, growth and action.* Amsterdam: Elsevier.

Cavalli-Sforza, L. (1997). Genes, peoples, and languages. *Proceedings of the National Academy of Sciences, 94*, 7719-7724.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*, 703-722.

Chinapah, V., H'ddigui, E. L. M., Kanjee, A., Falayajo, W., Fomba, C. O., Hamissou, O., Rafalimanana, A. & Byamugisha, A. (2000). *With Africa for Africa. Towards quality education for all. 1999 MLA project.* Pretoria: Human Sciences Research Council.

CIA (2012). *CIA (Central Intelligence Agency) World Factbook.* Retrieved January 13, 2011 from

www.cia.gov/library/publications/the-world-factbook/geos/mp.html.
Cliffordson, Ch., & Gustafsson, J.-E. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence*, *36*, 143-152.
Colom, R., Flores-Mendoza, C., & Abad, F. J. (2007). Generational changes on the Draw-a-Man Test: A comparison of Brazilian urban and rural children tested in 1930, 2002 and 2004. *Journal of Biosocial Science*, *39*, 79-89.
Conference des Ministres de l'Éducation des pays ayant le français en partage (2008). *Synthèse des résultats PASEC VII, VIII et IX.* Dakar. Retrieved from www.confemen.org/le-pasec/rapports-et-documents-pasec/synthese-des-resultats-pasec-vii-viii-et-ix.
Daley, T. C., Whaley, S. E., Sigman, M. D., Espinosa, M. P., & Neumann, Ch. (2003). IQ on the rise: The Flynn effect in rural Kenyan children. *Psychological Science*, *14*, 215-219.
Dawkins, R. (2008/1992). *The extended phenotype. The long reach of the gene.* Oxford: Oxford University Press.
Deary, I. J. (2012). Intelligence. *Annual Review of Psychology*, *63*, 453-482.
Diamond, J. (2011/2005). *Collapse: How societies choose to fail or succeed.* New York: Penguin.
Ducrest, A.-L., Keller, L., & Roulin, A. (2008). Pleiotropy in the melanocortin system, coloration and behavioural syndromes. *Trends in Ecology and Evolution*, *23*, 502-510.
Eherer, S. (2009). *Unterschiede in der Skala Arbeitsgedächtnis und Sprachverständnis bei normalsehenden, sehbehinderten und blinden Kindern.* [*Differences in working memory and comprehension between blind and normal-seeing children.*]. Graz: Unpublished Master's Thesis.
Farron, S. (2010). *The affirmative action hoax.* Quezon City: New Century Books.
Falch, T., & Massih, S. (2011). The effect of education on cognitive ability. *Economic Inquiry*, *49*, 838-856.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51.
Geiser, Ch., Lehmann, W., & Eid, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence, 36*, 556-563.
Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (II, pp. 945-1017). Amsterdam: North-Holland.
Gonzales, P. (2000). *Mathematics and science in the eighth grade.* Washington: National Center for Education Statistics.
Goodenough, F. (1926). *Measurement of intelligence by drawings*. New York: World Book Co.
Hansen, K. T., Heckman, J. J., & Mullen, K. J. (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics, 121*, 39-98.
Hanushek, E. A., & Woessmann, L. (2009). *Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation.* Bonn: IZA DP No. 4575.
Harrison, L. E. (2006). *The central liberal truth: How politics can change a culture and save it from itself.* New York: Oxford University Press.
Horn, W. (1983). *Leistungsprüfsystem LPS*. Göttingen: Hogrefe.
Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., Capelle, F. v., Paviot, L., & Vellien, J. (2010). *SACMEQ III Project Results: Student achievement levels in reading and mathematics*. Retrieved February 14, 2011 from www.sacmeq.org/downloads/sacmeqIII/WD01_SACMEQ_III_Results_Student_Achievement.pdf.
Hunt, E. (2011). *Human intelligence.* Cambridge: Cambridge University Press.
Jablonski, N. G., & Chaplin, G. (2000). The evolution of human skin coloration. *Journal of Human Evolution, 39*, 57-106.
Jensen, A. R. (2006). Comments on correlations of IQ with skin color and geographic-demographic variables. *Intelligence, 34*, 128-131.

Karabel, J. (2005). *The chosen: The hidden history of admission and exclusion at Harvard, Yale, and Princeton.* Boston: Houghton Mifflin.
Kaufman, A. S., & Weiss, L. G. (Eds.). (2010). The Flynn effect (special Issue). *Journal of Psychoeducational Assessment, 28*, 379-505.
Keller, H. (2010). *Kindliches Zeichnen im kulturellen Kontext.* [*Children's painting in cultural context.*] Osnabrück: Niedersächsisches Institut für frühkindliche Bildung und Entwicklung.
Khaleefa, O., Abdelwahid, S. B., Abdulradi, F., & Lynn, R. (2008). The increase of intelligence in Sudan 1964-2006. *Personality and Individual Differences*, *45*, 412-413.
Kratzmeier, H., & Horn, R. (1980). *Raven-Matrizen Test. Advanced Progressive Matrices (APM).* Weinheim: Beltz.
Luria, A. R. (1976/1974). *Cognitive development. Its cultural and social foundations*. Cambridge: Harvard University Press.
Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature*, *297*, 222-223.
Lynn, R. (2008). *The global bell curve. Race, IQ, and inequality worldwide.* Augusta: Washington Summit.
Lynn, R. (2009). Fluid intelligence but not vocabulary has increased in Britain, 1979-2008. *Intelligence*, *37*, 249-255.
Lynn, R. (2010). National IQs updated for 41 Nations. *Mankind Quarterly*, *50*, 275-296
Lynn, R., & Meisenberg, G. (2010a). The average IQ of sub-Saharan Africans: Comments on Wicherts, Dolan, and van der Maas. *Intelligence*, *38*, 21-29.
Lynn, R., & Meisenberg, G. (2010b). National IQs calculated and validated for 108 nations. *Intelligence*, *38*, 353-360.
Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations.* Westport: Praeger.
Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality.* Athens: Washington Summit.
Makuwa, D. (2010). *What are the levels and trends in reading and mathematics achievement*? SACMEQ Policy Issues Series, 2.

Retrieved February 14, 2011 from
www.sacmeq.org/downloads/policy/002-
SACMEQPolicyIssuesSeries-Studentachievement.pdf.
Meisenberg, G., Lawless, E., Lambert, E., & Newton, A. (2005). The Flynn effect in the Caribbean: Generational change of cognitive test performance in Dominica. *Mankind Quarterly, 46*, 29-69.
Meyer, J. W., Ramirez, F. O., & Soysal, Y. N. (1992). World expansion of mass education, 1870-1980. *Sociology of Education, 65*, 128-149.
Müller, A. (2002). *Raumvorstellung und logisches Denken bei Jugendlichen in Ostafrika.* [*Mental rotation and reasoning of teenagers in East Africa.*] Magdeburg: Unpublished Master's Thesis.
Petermann, F. & Petermann, U. (2007). *HAWIK-IV.* Bern: Huber.
Peters, M., Laeng, B., Lathan, K., Jackson, M., Zaiouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse Mental Rotations Test: Different versions and factors that affect performance. *Brain and Cognition, 28*, 39-58.
Raven, J. C. (1976/1943). *Advanced Progressive Matrices (APM).* Oxford: Oxford Psychologists Press.
Richter, L. M., Griesel, R. D., & Wortley, M. E. (1989). The Draw-a-Man test: A 50 year perspective on drawings done by black South African children. *South African Journal of Psychology, 19*, 1-5.
Rindermann, H. (2007). The *g*-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21,* 667-706.
Rindermann, H. (2008). Relevance of education and intelligence at the national level for the economic welfare of people. *Intelligence, 36*, 127-142.
Rindermann, H. (2011). Intelligenzwachstum in Kindheit und Jugend. [Increase of intelligence in childhood and youth.] *Psychologie in Erziehung und Unterricht, 58*, 210-224.

Rindermann, H., & Ceci, S. J. (2009). Educational policy and country outcomes in international cognitive competence studies. *Perspectives on Psychological Science, 4*, 551-577.
Rindermann, H., Sailer, M., & Thompson, J. (2009). The impact of smart fractions, cognitive ability of politicians and average competence of peoples on social development. *Talent Development and Excellence, 1*, 3-25.
Rindermann, H., & te Nijenhuis, J. (2012). Intelligence in Bali – A case study on estimating mean IQ for a population using various corrections based on theory and empirical findings. *Intelligence, 40*. doi: 10.1016/j.intell.2012.06.005.
Rindermann, H., Woodley, M. A., & Stratford, J. (2012). Haplogroups as evolutionary markers of cognitive ability. *Intelligence, 40*, 362-375.
Rushton, J. Ph. (2000/1995). *Race, evolution, and behavior. A life history perspective*. Port Huron: Charles Darwin Research Institute.
Rushton, J. Ph. (2004). Placing intelligence into an evolutionary framework or how g fits into the r-K matrix of life history traits including longevity. *Intelligence, 32*, 321-328.
Rushton, J. Ph., Skuy, M., & Fridjhon, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White engineering students in South Africa. *Intelligence, 31*, 123-137.
Segerstråle, U. (2000). *Defenders of the truth. The sociobiology debate*. New York: Oxford University Press.
Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjhon, P., & Rushton, J. Ph. (2002). Effects of mediated learning experience on Raven's matrices scores of African and non-African university students in South Africa. *Intelligence, 30*, 221-232.
Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist, 42*, 137-144.
Sowell, Th. (2004). *Affirmative action around the world: An empirical study*. New Haven: Yale University Press.

Statistics South Africa (2011). *Mid-year population estimates*. Pretoria: Statistics South Africa. Retrieved January 13, 2012 from www.statssa.gov.za/publications/P0302/P03022011.pdf.
Stelzl, I., Merz, F., Remer, H., & Ehlers, Th. (1995). The effect of schooling on the development of fluid and crystallized intelligence: A quasi-experimental study. *Intelligence, 21*, 279-296.
Stern, W. (1952/1914). *Psychologie der frühen Kindheit bis zum sechsten Lebensjahr.* [*Psychology of early childhood up to the sixth year of age.*] Heidelberg: Quelle & Meyer.
Studeny, P. (2008). Kognitive Fähigkeiten blinder und sehbehinderter Kinder im Ländervergleich Österreich und Südafrika. [*Cognitive abilities of blind and normal-seeing children in Austria and South Africa.*] Graz: Unpublished Master's Thesis.
Templer, D. I., & Arikawa, H. (2006). Temperature, skin color, per capita income, and IQ. *Intelligence, 34*, 121-128.
te Nijenhuis, J., Murphy, R., & Eeden, R. v. (2011). The Flynn effect in South Africa. *Intelligence*, *39*, 456-467.
Tuddenham, R. D. (1948). Soldier intelligence in world wars 1 and 11. *American Psychologist, 3*, 54-56.
UNDP (2010). *Human development report 2010.* New York: United Nations.
US Department of Labor (2009). *2008 Findings of the worst forms of child labor.* Retrieved January 12, 2012 from www.dol.gov/ilab/programs/ocft/PDF/2008OCFTreport.pdf
Wainwright, M., Wright, M. J., Geffen, G. M., Geffen, L. B., Luciano, M., & Martin, N. G. (2005). The genetic basis of academic achievement on the Queensland Core Skills Test and its shared genetic variance with IQ. *Behavior Genetics*, *35*, 133-145.
Wicherts, J. M., Dolan, C. V., Carlson, J. S., & Maas, H. L. J. v. d. (2010a). Raven's test performance of sub-Saharan Africans: Average performance, psychometric properties, and the Flynn Effect. *Learning and Individual Differences*, *20*, 135-151.
Wicherts, J. M., Dolan, C. V., & Maas, H. L. J. v. d. (2010b). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence*, *38*, 1-20.

Winship, Ch., & Korenman, S. (1997). Does staying in school make you smarter? The effect of education on IQ in The Bell Curve. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.), *Intelligence, genes and success: Scientists respond to the Bell Curve* (pp. 215-234). New York: Springer.

Part 4: Personality and Development

Chapter 5

The General Factor of Personality: Substance or Artefact?

Paul Irwing
University of Manchester
Business School East, The University of Manchester, Booth Street West, Manchester, M15 6PB.

Abstract

While it is now widely recognized that a general factor (GFP) can be extracted from most personality data, this finding has been subject to numerous critiques: (1) that the GFP is an artefact due to socially desirable responding; (2) that it is factorially indeterminate; (3) that it can be more parsimoniously modelled using blended variables; (4) that it shows less genetic variance due to dominance than should be true of a fitness trait; (5) that it correlates more weakly with g than would be predicted from Life History theory; (6) that it cannot be recovered across personality inventories. We present new evidence and argument to show that each of these critiques is open to reasonable doubt.

1. Introduction

Arguably, Francis Galton (1887) was the first to propose that a general factor underlies measures of personality. In their review, Rushton and Irwing (2011) trace a line of early research on this topic beginning with Galton and progressing through the work of Pearson, Davenport, Webb, and Freud. Recently, it has been demonstrated that a general factor of personality can be located in most accepted measures of personality, be they normal or abnormal, and for the Five Factor Model, this finding has been replicated five times with samples ranging in size from 4,000 to 628,640 (Just, 2011, Rushton & Irwing, 2011). A qualification to this is that when modelling primary scale data using CFA, it is necessary to include cross-factor loadings and correlated errors, however, this is a problem which is generic to personality data (e.g. Hopwood & Donnellan, 2010). Exceptionally, there are studies which do not find a general factor, but they are rare (de Vries, 2011).

Despite the consistent emergence of a general factor of personality within self-reported data, there is now a considerable body of empirical data which raises significant questions as to the veridicality of the GFP. Here we consider six major critiques.

2. Social Desirability

The overwhelmingly dominant view of the GFP is that it represents an artefact due either to evaluative bias or responding in a socially desirable manner. The social desirability hypothesis probably dates from a series of brilliant studies by Allen Edwards beginning in 1953. Three representative studies from this series established that loadings on the first principal component of 60 and 48 MMPI scales, correlated at 0.98 and 0.995 respectively with these same scales' correlations with social desirability (Edwards, Diers & Walker, 1962; Edwards & Walsh, 1963), and that loadings on the first principal component extracted from personality adjectives

devised by Peabody correlated at 0.90 with their correlations with social desirability ratings (Edwards, 1969).

Subsequently Bäckström (2007) extended this work by showing that 100-IPIP items designed to measure the Big 5 administered to 2,019 participants provided a good fit to a Bi-factor model. The GFP from this data correlated at 0.98 with a latent variable of social desirability. In a second study, Bäckström, Björklund, & Larsson (2009) were able to show that by rewording the 100 IPIP items so that they were evaluatively neutral, loadings on the GFP were reduced from an average of 0.56 to 0.09. Most recently, in a sophisticated study making use of EFA and the recently developed technique of exploratory structural equation modelling, Pettersson, Turkheimer, Horn and Menatti (2011) analyzed 120 items inspired by the Peabody adjectives originally used by Edwards (1969), which were administered to 619 participants. They were able to show that loadings on the GFP derived from this data correlated at 0.86 with mean social desirability ratings. Findings such as these have lead most researchers to conclude that general factors extracted from personality data are probably an artefact due to some form of evaluative bias.

The interpretation that the GFP is an artefact of socially desirable responding has been bolstered by a series of Multitrait-Multimethod (MTMM) studies. The basic logic of MTMM studies is that if a correlation between two traits is due to a method specific artefact, then if the same traits are measured using different methods then the correlation between them should either be substantially reduced or tend to zero. There are a large range of MTMM models (e.g. Widaman, 1985; Eid, Nussbeck, Geiser, Cole, Gollwitzer, & Lischetzke, 2008) and the various studies which have investigated this question have used a variety of these. By far and away the most common finding from these studies is that the GFP is very substantially reduced in magnitude when examined using MTMM methodologies based on structural equation modelling (Anusic, Schimmack, Pinkus, & Lockwood, 2009; Biesanz & West, 2004; Chang, Connelly, & Geeza, 2012; Danay & Ziegler, 2011;

DeYoung, 2006; Riemann & Kandler, 2010). Exceptions to this finding include those derived from a relatively simple MTMM model in which self-, teacher- and parent-ratings were treated as equivalent indicators (Rushton et al., 2009), or from correlations between GFPs derived from different ratings or personality inventories (Loehlin, 2011; Veselka, Just, Jang, Johnson, & Vernon, 2012; Zawadzki & Strelau, 2010). Certainly the latter are open to the criticism that the observed correlations may be due to variance attributable to sub-factors rather than the GFP (Keith, Reynolds, Patel & Ridley, 2008).

To summarize, the general factor of personality correlates so highly with measures of social desirability as to be almost indistinguishable from it, and is considerably reduced in magnitude when measured across raters in MTMM analyses. Altogether then, the extant evidence appears overwhelmingly to support the consensus view that, while the GFP may be extracted from mono-method data, it represents no more than an evaluative bias. On the basis of this evidence, the social desirability hypothesis appears to be inescapably correct, yet there are aspects of it which seem implausible. With regard to Edward's studies of the MMPI, there are countless studies which show that the MMPI, as it was designed to do, measures psychopathology (Hiller, Rosenthal, Bornstein, Berry, & Brunell-Neuleib, 1999). So it might be expected that the first principal component of the MMPI, accounting for 43% of the reliable variance (Edwards et al., 1962) measures psychopathology. Instead, it is perhaps not hard to believe that individuals evidence a bias towards socially desirable responding, but if socially desirable responding is to explain the GFP, that would require systematic variability in this response set. That is, at one extreme individuals should claim desirable traits they do not possess, but equally for this explanation to be correct, it would be required that some individuals evidence a genetic propensity (see below) to admit to pathological symptoms they do not possess. The question is why, and why should such an individual difference form such a large part of the variability in response to personality items?

It has to be admitted that evidence for the straightforward explanation that the first principal component of the MMPI measures psychopathology, is neither as abundant, nor of as good quality as that provided by Edwards in support of the social desirability hypothesis. Nevertheless it exists. Crowne and Marlowe (1960), and Heilbrun (1964) have argued that Edward's Social Desirability scale (SDS) is highly confounded with psychopathology. Edward's SDS was constructed by taking items from the MMPI which 10 judges unanimously rated as socially desirable. For such a consensus to exist, these items must be statistically deviant, and deviant items in a measure of psychopathology are likely to measure the extremes of this. Heilbrun (1964) provided two tests of the thesis that social desirability rankings and psychopathology are confounded. In one, 15 personality variables were rated for psychological adjustment by 25 psychologists with doctorates, and the rank order correlations calculated for these rankings with social desirability ranks derived from Edwards, and personal desirability rankings by college students using Edwards' methodology. These correlations were respectively 0.78 and 0.82. In a second study, Heilbrun reports social desirability rankings of 10 MMPI scales, which were correlated with Point-Biserial rs derived by correlating scores with membership in either a group of normal subjects (N=900) versus psychopathic hospital patients (N=100), or membership of a group of normal college students (N=270) versus maladjusted counselling service clients (N=30). The respective rank order correlations were 0.75 and 0.77. Given that rank order correlations are attenuated, these studies show a high degree of confounding between social desirability rankings and psychopathology. Ultimately, therefore, all the evidence pointing to the GFP in mono-method data as being attributable to a social desirability artefact is equally supportive of the interpretation that it is a measure of psychopathology.

In this context the MTMM data appears crucial. One of the underlying assumptions of MTMM is that correlations between traits on a single method can be biased by artefacts or method bias, whereas correlations across methods will be less susceptible to such

effects (Eid, Lischetzke, Nussbeck & Trierweiler, 2003). Therefore, if higher order factors of personality are the result of method bias and/or artefacts, theoretically they should not emerge from cross method correlation matrices. As noted above, evidence from MTMM studies has suggested that the GFP has largely failed to emerge in cross method data analyses and is therefore due to some form of method bias.

However, while it is generally concluded that failures of the GFP to emerge across raters in MTMM analyses are because it constitutes an artefact, there are other possible reasons. Elsewhere, it has been argued that there is considerable evidence for the situational specificity of human behaviour (Bandura, 1997, Mischel & Shoda, 1995). In consequence, it may be that the biggest component of other ratings is situational specificity. Many previous researchers have suggested this, for example McCrae et al. (2008) noted that, "Although personality psychologists usually interpret agreement as evidence of accuracy and disagreement as evidence of method bias, neither of these is necessarily the case: Agreement may be false consensus, and disagreement may reflect unique knowledge" (pp. 452).

The empirical evidence which exists strongly suggests that the largest component of most ratings comprises unique knowledge. Diary studies of personality, are perhaps most informative. These typically require respondents to complete abbreviated measures repeatedly over a period of between one and three weeks, each time rating an immediately prior sequence of behaviour of between five minutes and three hours duration. The findings of these studies consistently demonstrate that, while there is a high degree of consistency in mean levels, intra-individual variability in personality is greater than inter-individual variability. Fleeson and Gallagher (2009) provide estimates of the percentage of intra-individual variability, based on 21,871 reports, at 78%, 63%, 75%, 62% and 49% for Extraversion, Agreeableness, Conscientiousness, Emotional Stability and Intellect respectively. It has been shown that personality expression varies depending on status relationships, intimacy, the overall context (e.g. work versus

home), whether approach or avoidance goals are operative, and the cultural context (Heller, Komar, & Lee, 2007; Moskowitz, 2009; Wood & Roberts, 2006). Because each individual is exposed to a systematically different slice of the focal individual's behaviour, it follows that each rater largely contributes unique knowledge. Moreover, although it was not their focus, a meta-analysis by Connelly and Ones (2010) provides strong support for the contention that observer ratings are largely rater specific. For Big Five traits they found inter-rater reliabilities to range from 0.32 to 0.43 on average. Moreover, when these were corrected for test-retest reliabilities, they still ranged only from 0.39-0.51. These corrected reliabilities reflect the population overlap between ratings, which is clearly only 16-25%, This confirms that most variance in ratings is unique to the individual rater. Under this circumstance the correct way to combine personality ratings is additively in order to form composites (Bollen & Bauldry, 2011). Given these findings, to model ratings in terms of their covariance, as do all current MTMM models, is to eliminate the majority of valid variance, and hence to seriously distort measurement.

In short, evidence to date which is supportive of the social desirability hypothesis is based on inference, and in the case of MTMM analyses, the very complex assumptions underlying these inferences may not be correct. For these reasons, it would be desirable to make a more direct test of the social desirability hypothesis. In fact, Irwing, Rushton and Booth (2011) have already reported such a test which is summarized here. It is based on the long established contention that forced-choice item formats for the assessment of personality provide a higher degree of control over various forms of response bias including impression management and halo, than do Likert scaled items (Cheung & Chan, 2002; Christiansen, Burns & Montgomery, 2005; Jackson & Wroblewski, 2000; Martin, Bowen, & Hunt, 2002; Saville & Wilson, 1991). The principle underlying forced choice response formats is that respondents are asked to choose between blocks of items equated for social desirability, such that choosing one socially desirable response precludes the choice of other equally socially desirable

responses, thus effectively controlling out evaluative bias, at least to the extent that it is possible to equate items.

One personality test, the OPQ, which measures personality in a work specific context, exists in both a normative (OPQ32n), and forced-choice format (OPQ32r). That the forced choice version of the OPQ reduces effects due to both halo and socially desirable responding has both been hypothesized and demonstrated empirically (Baron, 1996; Bartram, 1996, 2007; Martin, et al., 2002; Saville & Wilson, 1991). It has also been shown more generally that forced choice response formats reduce faking (e.g. Christiansen et al., 2005).

Given the evidence that forced choice formats effectively control for halo and socially desirable responding both generally and specifically as applied to the OPQ, it follows that comparison of the higher-order factor structures of the OPQ in forced-choice and normative forms should provide a direct test as to whether these higher-order factors are due to halo or evaluative biases. If the higher-order factors are due to halo or evaluative biases, then they should either be greatly reduced in magnitude or disappear completely when using forced-choice measurements.

However, it has generally been concluded that ipsative data are not suitable for common data analysis, especially factor analysis, because the sum of variables equals a constant for all respondents (Chan & Bentler, 1993; Cheung, 2004; Cornwell & Dunlap, 1994; Meade, 2004). However, Brown and Maydeu-Olivares (2011) have shown that using a two-dimensional IRT model (the Thurstonian IRT model), it is possible to provide precise measurement of underlying personality traits with none of the problems typically associated with ipsative data. Brown and Bartram (2009) have applied this method of estimation to the Occupational Personality Questionnaire (OPQ32r). In consequence, it is possible to directly test the halo/social desirability hypothesis by comparing the higher-order factor structures derived from the normative version of the OPQ with the forced-choice version scored using the Thurstonian IRT Model.

2.1. Method

2.1.1 Samples

Data were available in the form of correlations between the 32 primary scales of the OPQ, for the UK standardization sample of the OPQ32n (N = 2,028) and for a largely student calibration sample of the OPQ32r (N = 518).

2.1.2 Measures

The OPQ32 has been translated into 30 languages and is used worldwide in the selection of managers/professionals and for counselling and development in professional groups. It measures 32 personality characteristics indicating people's preferred or typical style of work behavior. In the technical manual, the 32 scales are typically grouped into three topical areas: *Relationship with People* (the first 10), *Thinking Style* (the next 12), and *Feelings and Emotions* (the final 10), and can be joined by a potential fourth, *Dynamism*, composed of scales such as Vigorous, Achieving and Competitive (Bartram, Brown, Fleck, Inceoglu, & Ward, 2006). The 32 scales together with Cronbach's alpha reliabilities for the OPQ32n, followed by IRT composite reliabilities for the OPQ32r (both in brackets) are: Persuasive (0.76, 0.83), Controlling (0.84, 0.91), Outspoken (0.76, 0.86), Independent Minded (0.70, 0.77), Outgoing (0.84, 0.89), Affiliative (0.81, 0.84), Socially Confident (0.85, 0,87), Modest (0.84, 0.81) , Democratic (0.65, 0.74), Caring (0.72, 0.81), Data Rational (0.80, 0.88), Evaluative (0.70, 0.80), Behavioral 0.84, 0.79), Conventional (0.74, 0.68), Conceptual (0.78, 0.78), Innovative 0.84, 0.89), Variety Seeking 0.70, 0.77), Adaptable (0.73, 0.87), Forward Thinking (0.78, 0.87), Detail Conscious (0.76, 0.89), Conscientious (0.74, 0.84), Rule Following (0.84, 0.89), Relaxed (0.86, 0.87), Worrying (0.86, 0.78), Tough Minded (0.0.87, 0.80), Optimistic (0.83, 0.81), Trusting (0.84, 0.88), Emotionally Controlled (0.79, 0.86), Vigorous (0.79, 0.88), Competitive (0.77, 0.87), Achieving (0.81, 0.79), and Decisive (0.76, 0.83).

There is no consensus as to the most appropriate factor structure for the OPQ. Matthews, Stanton, Graham, and Brimelow (1990) found that a five factor solution explained a majority of the variance in 30 OPQ scales, as did Ferguson, Payne and Anderson (1994) using the 19 scales of the OPQ FMX5-Student Version. Matthews and Stanton (1994) found that a six factor solution was the most consistent across several samples with the factors representing an Activity factor plus the Big Five (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). They also suggested the six factors could be reduced to Eysenck and Eysenck's (1975) Big Three (Extraversion, Neuroticism, and Psychoticism). More recently, it has been reported in the Technical Manual of the OPQ, that a scree test applied to two large population samples (Ns = 2,028, 1,053) and two managerial samples (Ns = 2,009, 644), resulted in six factors being consistently identified (Bartram et al., 2006).

2.1.3 Strategy of Analysis

The data was analyzed in six stages. First, using the forced-choice sample, an exploratory factor analysis of the correlations between the 32 primary scales was carried out. Second, this solution formed the basis of a first-order confirmatory factor model, also fitted to the forced-choice data. Third, a second-order was fitted combining the first-order factor structure with a GFP at the apex, again using the forced choice data. Fourth a GFP was fitted to the correlations between the six first-order factors (see Table 3 and Figure 1). Fifth, the primary Confirmatory Factor Analysis (CFA) model combined with the GFP (see Table 2), was cross validated on the correlations between the 32 scales in the normative data. Sixth, an invariance test of the GFP as shown in Figure 1 was conducted simultaneously in both samples. In short, first an essentially exploratory analysis was carried out using the forced-choice data, which gave rise to CFA models. These were then cross validated on the correlations provided by the normative sample. This sequence of analyses approximates the ideal strategy outlined for model testing by Jöreskog (1993). Designated "strictly confirmatory," prior theory

and research point to the correctness of a single model, which is then tested in a representative sample and, if confirmed, shows that the model is generalizable.

2.2. Results

An exploratory factor analysis was conducted in the forced-choice sample, using Mplus with maximum likelihood estimation and a Promax rotation, testing for a six factor solution as suggested in the manual (Brown & Bartram, 2009). This solution differed somewhat from that of Brown and Bartram (2009). This is partly because, since the concern was with the higher-order factor structure of the OPQ32n, an oblique rather than an orthogonal rotation was employed in order to correctly capture the correlations between the six first-order factors. Nevertheless the factors which emerged were not hugely different from those reported by Brown and Bartram (2009), and were labelled: (1) Openness (Unconventionality); (2) Conscientiousness; (3) Extraversion; (4) Leadership; (5) Openness (critical thinking); and (6) Emotional Stability. The nature of factor 4 differs somewhat from the Agreeableness factor found by Brown and Bartram (2009), and perhaps reflects the work oriented nature of the OPQ.

Table 1. Fit statistics for confirmatory factor models fitted to the OPQ32r and OPQ32n.

Model	x^2	*df*	NNFI	RMSEA	SRMR
Opq32r					
1. Six correlated factors	1546.7	411	.90	.07	.058
2. Six uncorrelated factors	2173.2	427	.85	.087	.15
3. Model 2 plus GFP	1691.7	422	.89	.072	.078
4. GFP	138.5	9	.82	.16	.081
5. GFP with correlated errors	0.2	7	1.01	.000	.000
6. Preferred model	1624.7	420	.89	.071	.071
Cross validation and invariance on OPQ32n					
7. Validation of Model 6	4006.7	422	.90	.063	.061
8. Strictly invariant GFP	64.5	28	.99	.035	.059/.016

This six factor solution was then directly entered into a confirmatory factor analysis. Initially only loadings above .30 were included in the model, but a further 17 loadings and 12 correlated errors, each corresponding to the largest modification index, were sequentially added until the model provided an acceptable fit (see Table 1, Model 1 and Table 2). In order to assess whether a higher-order factor structure was plausible, next a model was fitted in which the six first-order factors were uncorrelated. This model provided a very poor fit to the data, supporting the presence of higher-order factors (see Table1, Model 2).). Next a GFP was added, at the apex, to account for the correlations between the first order factors, which model fit reasonably well (see Table 1, Model 3).

Table 2. First-order common factor structure of the OPQ32r and (after backslash) OPQ32n.

	Factors						Error
	1	2	3	4	5	6	
1. Persuasive	.57/.41		.22/.29				.52/.63
2. Controlling		.47/.44	.40/.35	.23/.27			.42/.45
3. Outspoken			.60/.60	.37/.42			.56/.54
4. Independent minded			.31/.29	.60/.55			.59/.67
5. Outgoing			.81/.75				.34/.44
6. Affiliative			.49/.48	-.44/ -.40			.51/.55
7. Socially confident			.44/.30			.52/.53	.44/.53
8. Modest			-.61/ -.54				.63/.71
9. Democratic				-.74/ -.63	.38/ .37		.46/.58
10. Caring		.29/.33		-.57/ -.56			.69/.67
11. Data rational		.31/.36					.90/.87
12. Evaluative		.31/.30			.58/.46		.43/.55
13. Behavioural				-.28/ -.33	.62/.69		.63/.52
14. Conventional	-.85/ -.73	.24/.22					.47/.65

	Factors						Error
	1	2	3	4	5	6	
15. Conceptual					.78/.70		.39/.50
16. Innovative	.78/.69						.39/.52
17. Variety seeking	.47/.53						.78/.72
18. Adaptable	.36/.41					-.41/ -.40	.85/.86
19. Forward thinking		.59/.53					.66/.71
20. Detail conscious	-.50/ -.60	.91/.92					.48/.61
21. Conscientious	-.50/ -.58	.92/.91					.47/.62
22. Rule following	-.90/ -.87	.58/.69					.50/.66
23. Relaxed		-.41/ -.22				.86/.67	.46/.68
24. Worrying					-.21/ .02	-.73/ -.72	.33/.50
25. Tough minded			-.22/ -.09			.64/.54	.60/.73
26. Optimistic				-.25/ -.32		.57/.60	.68/.62
27. Trusting				-.52/ -.51		.42/.36	.65/.68
28. Emotionally controlled			-.61/ -.51				.62/.74
29. Vigorous		.63/.50					.60/.75
30. Competitive			.28/.13	.45/.40			.75/.84
31. Achieving	.22/.35	.67/.42					.32/.47
32. Decisive			.32/.30	.46/.38		.21/.19	.60/.71
GFP	.99/.99	.63/.75	.44/.53	.46/.34	.62/.69	.50/.57	

Note. 1 = Openness (unconventionality), 2 = Conscientiousness, 3 = Extraversion, 4 = Leadership, 5 = Openness (critical thinking), 6 = Emotional Stability

However, well-fitting models as evidenced by global indices may nevertheless evidence localized misfit (Tomarken & Waller, 2003), so the fit of the second-order GFP to the correlations between the six broad factors was directly tested (see Table 3). This model provided a poor fit to the data (see Table 1, Model 4). However, with the addition of two correlated errors (see Figure 1),

the fit was excellent (see Table1, Model 5). In the light of these findings, our final preferred model of the correlations between the 32 scales of the OPQ32r included six primary factors, a GFP, and two correlated residuals. This model evidenced a moderate level of fit (see Table 1, Model 6).

This preferred model of the OPQ32r data, incorporating both six first-order factors and a GFP at its apex, was cross validated onto the correlations between the OPQ scales derived from the normative sample. This model evidenced good fit according to the SRMR and RMSEA and showed an improved level of fit compared with the same model fitted to the forced-choice data (see Table 1, Model 7, and Table 2). Twenty six of the 29 parameters, which were included in the factor model based on the magnitude of their modification indices, were significant in the normative sample, demonstrating a good level of generalizability. However, the crucial issue from the perspective of our research hypothesis was that the GFPs extracted from the forced-choice and normative samples should be more or less equivalent. A formal test of the similarity of the factor structures of the GFP under conditions of normative and forced-choice measurement was carried out using multi-group confirmatory factor analysis. A GFP was fitted to the correlations between the six first-order factors, in both the ipsative and normative samples simultaneously, specifying that all parameters were strictly invariant. The fit of this model was excellent (see Table 1, Model 8). The GFPs fitted to the forced-choice and normative data are therefore effectively, identical. Figure 1 presents the parameter estimates for the GFP showing the common metric fully standardized solution. The undesirably high loading of Unconventionality on the GFP represents a Heywood case. Unfortunately, this is a not uncommon problem in factor analysis (MacCallum, Widaman, Zhang, & Hong, 1999).

Table 3. Correlations between the six first-order factors of the OPQ32r and OPQ32n (in brackets).

	1	2	3	4	5	6
1	1.00					
2	.62(.74)	1.00				
3	.43(.52)	.27(.40)	1.00			
4	.45(.33)	.28(.25)	-.13(-.16)	1.00		
5	.61(.68)	.39(.52)	.27(.36)	.28(.23)	1.00	
6	.49(.56)	.51(.61)	.22(.30)	.23(.19)	.31(.39)	1.00

Note. 1 = Openness (unconventionality), 2 = Conscientiousness, 3 = Extraversion, 4 = Leadership, 5 = Openness (critical thinking), 6 = Emotional Stability

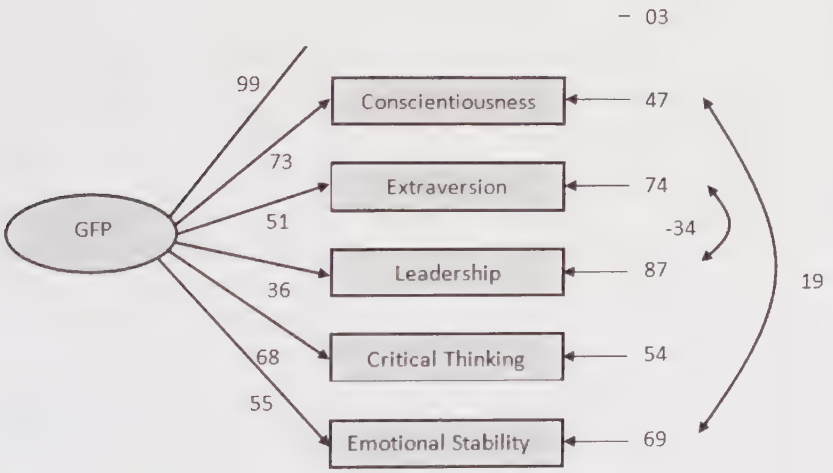


Figure 1. Second-order factor structure of the OPQ32r and OPQ32n: The common metric. completely standardized solution.

The most plausible interpretation of these results is that the GFP is not due to artefacts such as halo effects and social

desirability responding because a highly similar GFP emerged from both the normative and forced-choice versions of the Occupational Personality Questionnaire. The forced-choice items had been matched for social desirability and so substantially reduced the possibilities of artefact (Christiansen, et al., 2005; Jackson & Wroblewski, 2000; Martin et al., 2002). The forced-choice format has also been shown to successfully reduce uniform response biases such as halo and acquiescence (Cheung & Chan, 2002). It follows, therefore, that if response biases such as social desirability and halo effects had produced an artefactual GFP in the normative measures, this would have been greatly reduced or absent when using forced-choice measures. In short, the covariance structure of normative measures should have been substantially biased, whereas it should have been highly resistant to bias in forced-choice measures (Brown, 2008). Much as expected, the covariance structures across normative and forced-choice versions of the OPQ32 were different, but not at much above chance levels, as evidenced by the excellent fit shown by the multi-group analysis. This appears to rule out the possibility that a GFP recovered from population representative samples is due to an artefact of response bias. A limitation of this finding is that it was necessary to include two correlated errors for the GFP to show excellent fit. However, second-order factor structures of personality inventories are commonly problematic.

3. Conditions for the valid measurement of the GFP

Here we conceive of the general factor of personality as a unitary dimension ranging from the adaptive to the psychopathological, which corresponds both with the early ideas of Davenport (1911) and with a growing body of research evidence which suggests that normal and clinical personality should be measured on a single continuum (e.g. Markon, Krueger & Watson, 2005; Samuel, Simms, Clark, Livesley & Widiger, 2010). By adaptive we mean the extent to which ones' personality is adapted to situational requirements, in order to attain desired goals.

If a GFP conforming to this definition exists, it should be manifest in all measures of personality with adequate measurement properties. However, there is clearly substantial variability across inventories examined to date as to the percentage of variance accounted for by the GFP. The question therefore arises as to what are the required measurement properties in order to provide a consistent and unbiased estimate of the GFP. Clearly, in conformity with the definition, the inventory should be comprised of items reflecting the full range from normal to abnormal. A second requirement is that scales must be unidimensional. Unidimensionality is a necessary condition in order to obtain consistent and unbiased estimates of the correlations between the primary scales of personality (Hattie, 1985). There are further questions as to who should supply ratings and how they should be combined. This is a matter for current research. To date it appears that multiple other raters who are closely acquainted with the subject may provide the most valid source of ratings (Connelly & Ones, 2010). We have suggested above that this data may be best represented by composite scales formed by combining individual ratings additively. There is a further question as to how many dimensions of personality should be included. Currently, there are findings which suggest that the total number of personality facets is considerably higher than previously thought (Samuel et al., 2010; Booth, 2011). By analogy with research on g, according to current knowledge, it seems likely that a valid measure of the GFP should include all the broad factors of personality with at least four indicators per factor, estimated using hierarchical confirmatory factor analysis (Major, Johnson & Bouchard, 2011).

Unfortunately, to date there is little evidence that current measures of personality meet these stipulations (Hopwood & Donnellan; 2010; Pace & Brannick, 2010; Vassend & Skrondal, 2011). Some measures such as the 16PF do appear to be largely unidimensional, but fail to measure the full range of personality (Booth, 2011), while others such as the MMPI may contain measures of both normal and abnormal personality, but are probably not unidimensional. Nevertheless, it is probably the case

that the item pool for an inventory like the MMPI is more likely to provide an adequate basis for estimation of a GFP than are many other inventories.

Sampling is another factor which influences estimates. Large population representative samples are necessary, since samples which do not satisfy these conditions will be subject to range restriction which attenuates correlations, and sampling variability which biases estimates.

4. Factor saturation

The conditions necessary for the estimation of a valid GFP are relevant to a critique advanced by Revelle and Wilt (2009). An attendant problem of the common factor model is factor score indeterminacy (Mulaik, 2005). As a consequence of this, factors which account for less than 50% of the measured variance may be associated with factor scores which are negatively correlated, and therefore indeterminate. It has been argued that statements about the suitability of general factors should primarily be based on estimates of McDonald's omega hierarchical (ώh: Revelle & Zinbarg, 2009). Revelle and Wilt (2009) report that the average ώh estimate for published GFP studies is 0.38 compared to 0.73 for studies of cognitive abilities and thus maintain that the GFP is indeterminate, whereas g is not. In the context of this argument cognizance of the conditions required in order to obtain a consistent and unbiased estimate of the GFP are crucial. While recognizing its inadequacy, Irwing, Booth, Nyborg and Rushton (2012) argued that the MMPI administered to a population representative sample might provide an indication of the factor saturation of the GFP. In this sample, the GFP attained a McDonald's ώh of 0.75, higher than that of g. Some caution is required since substantial item overlap biases the correlations between the MMPI scales (Helmes & Reddon, 1993), but nevertheless this finding suggests that, properly measured, the GFP may not be indeterminate.

5. Blended variable models

In 2009, Ashton, Lee, Goldberg, and de Vries offered an alternative explanation of the observed correlations between Big Five scales. They suggested that same sign blends may produce cross scale loadings. If these exist and are not modelled, they would give rise to spurious correlations amongst Big Five scales. If true this could possibly explain the finding of higher order factors amongst Big Five scales. However, the majority of evidence subsequent to 2009 has involved analysis of primary scale level data or even items rather than Big Five scales. Also in the numerous SEM analyses of these data, unmodelled cross factor loadings would have lead to misfit, yet all the models presented fit. In consequence, blended variable models cannot explain the majority of evidence supportive of the GFP.

Beyond this there are question marks concerning the arguments and evidence presented. Perhaps inadvertently, in the data analyzed, the authors compared a higher order model specified a priori with a blended variable model in which they had optimized fit by freeing parameters indicated by modification indices to be the cause of misfit. Since they were, therefore, comparing a theoretically pre-specified model with one directly fitted to the data, it is perhaps not surprising that the latter provided a superior fit. In a second series of analyses the authors used data from the HEXACO, and found no evidence supportive of a GFP, which finding is consistent with de Vries' (2011) analysis of data from the HEXACO. The HEXACO, as compared with all other personality inventories produces atypically small correlations between its six broad factors. It is sometimes argued that any example of data in which there is no higher-order factor of personality disproves the existence of a GFP (de Vries, 2011). This is not the case. There are many reasons why personality scales may not correlate. Probably the most important is the phenomenon of rotational indeterminacy. Factor axes can in principle be rotated to any position in factor space. Some of these positions are orthogonal and some oblique, but they are all mathematically equivalent (Mulaik, 2005).

Irrespective of their true position, therefore, it is perfectly possible to place factor axes into a position such that scale scores do not correlate. This could quite probably explain the small correlations observed between the broad factors of the HEXACO, although there are many other possible explanations, some of them considered in section 2 above. Moreover, not all studies have failed to find a GFP when using data from the HEXACO. Veselka et al. (2009) did recover a convincing GFP from the HEXACO, though they did so using principal components analysis which tends to overestimate the magnitude of factor loadings (Widaman, 1993).

6. Behaviour genetic evidence

Behaviour genetic evidence is relevant to the GFP hypothesis for two reasons. Firstly, under Rushton's interpretation the GFP should show substantial genetic determination, whereas if the GFP is merely a psychometric artefact then such a prediction makes little sense. Secondly, if the GFP leads to greater reproductive success as suggested by Rushton and Irwing (2011) then it should have been under recent directional selection. From this it would be predicted that a substantial proportion of its genetic variation should be non-additive. This is because both theoretically and in practice, strong directional and to some degree stabilizing selection usually primarily erode additive genetic variance while not affecting dominance variance. Consequently, traits closely associated with fitness should exhibit high levels of dominance variance (Crnokrak, & Roff, 1995). Of the eight twin samples which have to date investigated the behaviour genetics of the GFP, all have found evidence of substantial genetic heritability and six of these studies have supported the presence of non-additive effects (Loehlin & Martin, 2011; Veselka, Schermer, Petrides & Vernon, 2009, Rushton et al., 2009; Rushton, Bons & Hur, 2008.

Moreover, a recent study of the behaviour genetics of cognitive abilities suggests that the proportion of variance attributed to non-additive genetic variation is consistently under-estimated. For general intelligence in adulthood, it is commonly

concluded that heritability is in the range from 75 to 85% and that all of this is additive in form. But there is considerable evidence for assortative mating for intelligence, and yet the commonly applied behaviour genetic models do not take account of this. In a study using an extended twin-family design, Vinkhuyzen, van der Sluis, Maes and Posthuma (2012) have shown using a phenotypic assortment model that variance in adult intelligence was due to non-shared environmental (18%), additive genetic factors (44%), phenotypic assortment (11%), and most crucially non-additive genetic factors (27%). If Rushton and Irwing's characterization of the GFP is correct then current studies will have similarly underestimated the non-additive genetic component in this.

7. Other critiques

There are a number of issues which this review can only consider briefly.

Firstly, there have been four studies of the correlation between the GFP and g which estimate this at between 0.23 and 0.28. Although there have been suggestions that this is incompatible with Rushton's Life History perspective, there are in fact many possible explanations for this positive but somewhat small correlation. Some of these have been outlined in Irwing et al., (2012). An additional consideration is that when a trait is clearly adaptive it is generally contended that directional selection will act to elevate mean levels of that trait and reduce to zero any variance on that trait in the population, unless there is a countervailing force such as mutation load or balancing selection. In any given population, it is perfectly possible that a slow life history strategy is largely advantageous, whereas a fast life history strategy is largely not. Under these circumstances, from a Life History perspective the correlation between the GFP and g should be small, as it is observed to be. However, as noted above, ratings of personality from a single source are highly unreliable so it may simply be that all current studies have underestimated the magnitude of the correlation between g and the GFP.

Secondly, there have been a number of studies evaluating the evidence for a GFP across inventories. This represents an MTMM approach using inventories rather than raters as the alternative method. In the first of these studies Irwing (2009), using confirmatory factor analysis applied to data from the Eugene Springfield community sample, found a model which evidenced acceptable fit (χsb2 = -1893.6; df = 864RMSEA = .088; CFI = 1.00; TLI = 1.09) and showed that the GFP from the JPI, the HPI, Saucier's mini-markers, and the BFI correlated from 0.71 to 1.02 (see Figure 2). In other words they were sometimes identical and sometimes just similar. Given the problems associated with personality measurement, and our stipulations as to the number of indicators required to validly estimate the GFP, these findings are largely supportive of the GFP, particularly since the correlations which deviate most from unity involve the inventories with the fewest indicators, which findings are very similar to those obtained in equivalent investigations applied to the invariance of g (Johnson, te Nijenhuis & Bouchard, 2008). Subsequently, Hopwood, Wright and Donnellan (2011) using seven inventories administered to the Eugene-Springfield community sample came to the opposite conclusion. They reached this conclusion despite the fact that their confirmatory factor model did show a GFP (Figure 2, p. 474). They cite estimation difficulties and poor fit as the justification for their conclusion. However, poor fit is more likely due to the fact they did not apply an MTMM analysis to MTMM data, and the estimation difficulties may be due to the same source, or alternatively the sheer complexity of their model in what is a relatively small data set. Other studies have also shown that GFPs derived from different inventories correlate substantially (Loehlin, 2011; Veselka et al., 2012; Zawadzki & Strelau, 2010).

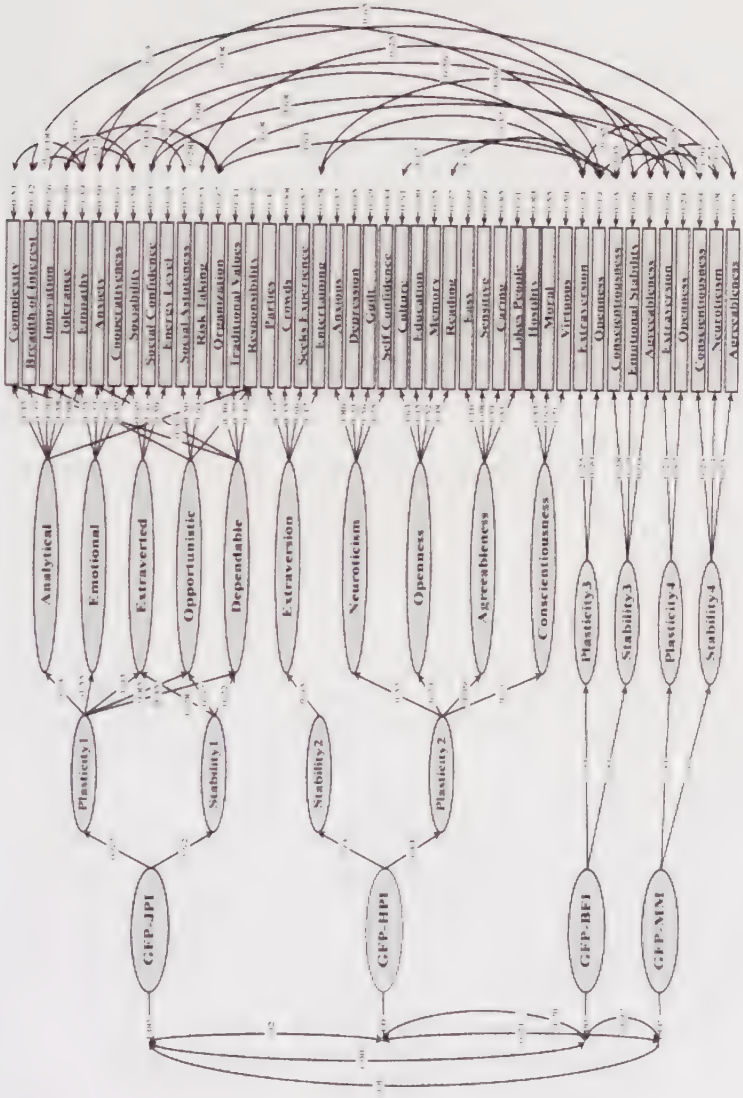


Figure 2. Hierarchical factor structure of four inventories (JPI, HPI, Mini-Markers, BFI) showing the correlations between the GFPs.

8. Conclusion

Given what is known about the problems of measurement in personality data, no theory of personality will be supported unambiguously. Unsurprisingly, therefore, evidence for the GFP is not unambiguous. Nevertheless, having reviewed six common critiques of the GFP, it has been shown that there is reasonable doubt attached to each. Therefore, given the potential importance of this construct a number of further lines of research appear justified.

First, it seems unlikely that the presence or otherwise of the GFP will be definitively established until a correct structure of personality has been derived. Unfortunately, this challenge is unlikely to be accomplished in the short term. In the meantime, perhaps an item level factor analysis of the MMPI might provide a moderately satisfactory measure of the GFP, for immediate use. There are of course current measures of the GFP, but none of these correspond to the GFP as conceived here (e.g. Caselles, Micó, & Amigó, 2011).

Second, there are two problems with current use of MTMM models. First, with few exceptions they do not take account of the non-equivalence of raters (Eid et al., 2008). Second, and perhaps more crucially, it is possible that personality ratings are more correctly modelled as composites rather than common factors. This possibility should be explored.

Third, GFPs derived from Likert and forced- choice formats have been shown to be effectively identical, which is incompatible with the social desirability hypothesis. Nevertheless, the GFP in this data required two correlated errors in order to achieve fit. It therefore remains to be shown that, when correctly measured personality data show a perfect positive manifold.

Fourth, there needs to be more extensive investigation as to whether valid measures of the GFP are factorially determinate.

Fifth, there is a need for behaviour genetic studies of personality which take account of assortative mating.

Sixth, genetically informed designs should be employed in order to determine why the GFP and g are weakly correlated. Although, the possibility should also be explored that this apparently weak correlation may be attributable to considerable error of measurement in single-rater estimates of the GFP.

Seventh, the equivalence of GFPs extracted from different inventories should be investigated using appropriate MTMM models.

References

Anusic, I., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2009). The nature and structure of correlations among Big Five ratings: The halo-alpha-beta model. *Journal of Personality and Social Psychology, 97*, 1142-1156.

Ashton, M. C., Lee, K., Goldberg, L. R., & de Vries, R. E. (2009). Higher-order factors of personality: Do they exist? *Personality and Social Psychology Review, 13*, 79-91.

Bäckström, M. (2007). Higher-order factors in a five-factor personality inventory and its relation to social desirability. *European Journal of Psychological Assessment, 23*, 63-70.

Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality, 43*, 335-344.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Baron, H. (1996). Strengths and Limitations of Ipsative Measurement. *Journal of Occupational and Organizational Psychology, 69*, 49-56.

Bartram, D. (1996). The relationship between ipsatized and normative measures of personality. *Journal of Occupational Psychology, 69*, 25-39.

Bartram, D. (2007). Increasing validity with forced-choice criterion measurement formats. *International Journal of Selection and Assessment, 15*, 263-272.

Bartram, D., Brown, A., Fleck, S., Inceoglu, I., & Ward, K. (2006). *OPQ32 technical manual.* Thames Ditton, UK: SHL Group.

Biesanz, J. C., & West, S. G. (2004). Towards understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality, 72*, 845-876.

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods*, *16,* 265-284.

Booth, T. W. (2011*). A review of the structure of normal range personality.* Unpublished doctoral dissertation, University of Manchester.

Brown, A. (2008). The impact of questionnaire item format on the ability to "fake good". In Brown, A. (chair): *Exploring the use of ipsative measures in personnel selection.* Symposium presented at the 6th International Conference of the International Test Commission, Liverpool.

Brown, A., & Bartram, D. (2009). *Development and psychometric properties of 0PQ32r: Supplement to the OPQ232 Technical Manual.* Thames Ditton, UK: SHL Group Ltd.

Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. Educational and *Psychological Measurement*, *71*, 460-502.

Caselles, A., Micó, J. C., & Amigó, S. (2011). Dynamics of the General Factor of Personality in response to a single dose of caffeine. *The Spanish Journal of Psychology*, *14*, 675-692.

Chan, W., & Bentler, P. M. (1993). The covariance structure analysis of ipsative data. *Sociological Methods & Research*, *22,* 214-247.

Chang, L., Connelly, B. S., & Geeza, A. A. (2012). Separating method factors and higher order traits of the Big Five: A meta-analytic multitrait-multimethod approach. *Journal of Personality and Social Psychology, 102*, 408-426.

Cheung, M. W. L. (2004). A direct estimation method on analyzing ipsative data with Chan and Bentler's method. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 217-243.

Cheung, M. W. L., & Chan, W. (2002). Reducing uniform response bias with ipsative measurement in multiple-group confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 9,* 55-77.

Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance, 18*, 267-307.
Connelly, B. S., & Ones, D. S. (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136*, 1092-1122.
Cornwell, J. M., & Dunlap, W. P. (1994). On the questionable soundness of factoring ipsative data: A response to Saville & Wilson (1991). *Journal of Occupational and Organizational Psychology, 67*, 89-100.
Crnokrak, P., & Roff, D. A. (1995). Dominance variance: Associations with selection and fitness. *Heredity, 75*, 530-540.
Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349-354.
Danay, E., & Ziegler, M. (2011). Is there really a single factor of personality? A multirater approach to the apex of personality. *Journal of Research in Personality, 45*, 560-567.
Davenport, C. B. (1911). Heredity in relation to eugenics. New York: Holt.
De Vries, R. E, (2011). No evidence for a general factor of personality in the HEXACO personality inventory. *Journal of Research in Personality, 45*, 229-232.
DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91,* 1138-1151.
Edwards, A. L. (1969). Trait and evaluative consistency in self-description. *Educational and Psychological Measurement, 29*, 737-752.
Edwards, A. L., Diers, C. J., & Walker, J. N. (1962). Response sets and factor loadings on sixty-one personality scales. *Journal of Applied Psychology, 46,* 220-225.
Edwards, A. L., & Walsh, J. A. (1963). The relationship between the intensity of the social desirability keying of a scale and the correlation of the scale with Edwards' SD scale and the first factor loading of the scale. *Journal of Clinical Psychology, 19*, 200-203.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple indicator CT-C(M-1) model. *Psychological Methods, 8*, 38-60.

Eid, M., Nussbeck, F. W., Geiser, C., Cole, D. A., Gollwitzer, M., & Lischetze, T. (2008). Structural equation modeling of Multitrait-Multimethod data: Different models for different types of data. *Psychological Methods, 13*, 230-253.

Ferguson, E., Payne, T., & Anderson, N. (1994). Occupational Personality Assessment: Theory, Structure and Psychometrics of the OPQ FMX5-Student. *Personality and Individual Differences, 17*, 217-225.

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology, 97*, 1097-1114.

Galton, F. (1887). Good and bad temper in English families. *Fortnightly Review, 42*, 21-30.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9*, 49-78.

Heilbrun, A. B. (1964). Social learning theory, social desirability, and the MMPI. *Psychological Bulletin, 61*, 377-387.

Heller, D., Komar, J., & Lee, W. B. (2007). The dynamics of personality states, goals, and well-being. *Personality and Social Psychology Bulletin, 33*, 898-910.

Helmes, E., & Reddon, J.R (1993). A perspective on developments in assessing psychopathology: A critical review of the MMPI and MMPI-2. *Psychological Bulletin, 1*13, 453-471.

Hiller, J. B., Rosenthal, R., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (1999). A comparative meta-analysis of Rorschach and MMPI validity. *Psychological Assessment, 11*, 278-296.

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14*, 332-346.

Hopwood, C. J., Wright, A. G. C., & Donnellan, M. B. (2011). Evaluating the evidence for the general factor of personality across multiple inventories. *Journal of Research in Personality, 45*, 468-478.

Irwing, P., (2009). *Just One GFP: Consistent Results from Four Test Batteries.* Paper presented at the meeting for the International Society for the Study of Individual Differences, Evanston, IL, July.

Irwing, P., Booth, T., Nyborg, H., & Rushton, J. P. (2012). Are g and the General Factor of Personality (GFP) correlated? *Intelligence, 40*, 296-305.

Irwing, P., Rushton, J. P,, & Booth, T. (2011). *A General Factor of Personality in the Occupational Personality Questionnaire (OPQ32) in Two Validity Samples.* Paper presented at the International Society for the Study of Individual Differences 11th Annual Conference. London, UK.

Jackson, D. N., & Wroblewski, V. R. (2000). The impact of faking on employment tests: Does forced choice offer a solution? *Human Performance, 13,* 371-388.

Johnson, W., te Nijenhuis, J., & Bouchard, T. J. Jr. (2008). Still just 1 g: Consistent results from five test batteries. *Intelligence, 36,* 81-95.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294-316). London: Sage.

Just, C. (2011). A review of literature on the general factor of personality. *Personality and Individual Differences, 50,* 765-771.

Keith, T. Z., Reynolds, M. R., Patel, P. G., & Ridley, K. P. (2008). Sex differences in latent cognitive abilities ages 6 to 59: Evidence from the Woodcock Johnson III tests of cognitive abilities. *Intelligence, 36,* 502-525.

Loehlin, J. C. (2011). Correlation between general factors for personality and cognitive skills in the National Merit twin sample. *Journal of Research in Personality, 45*, 504-507.

Loehlin, J. C., & Martin, N. G. (2011). A general factor of personality: Questions and elaborations. *Journal of Research in Personality, 45,* 44-49. (See also Corrigendum, *45,* 258).

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4,* 84-99.
McCrae, R. R., Yamagata, S., Jang, K. L., Riemann, R., Ando, J., Ono, Y., et al. (2008). Substance and artefact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 95*, 442–455.
Major, J. T., Johnson, W., & Bouchard, T. J. Jr. (2011). The dependability fo the general factor of intelligence: Why small single=factor models do not adequately represent g. *Intelligence, 39,* 418-433.
Markon, K.E., Krueger, R.F., & Watson, D. (2005). Delineating the Structure of Normal and Abnormal Personality: An Integrative Hierarchical Approach. *Journal of Personality and Social Psychology, 88*, 139-157.
Martin, B. A., Bowen, C.-C., & Hunt, S. T. (2002). How effective are people at faking on personality questionnaires? *Personality and Individual Differences, 32,* 247-256.
Matthews, G., & Stanton, N. (1994). Item and Scale Factor Analyses of the Occupational Personality Questionnaire. *Personality and Individual Differences, 16,* 733-743.
Matthews, G., Stanton, N., Graham, N.C., & Brimelow, C. (1990). A Factor Analysis of the Scales of the Occupational Personality Questionnaire. *Personality and Individual Differences, 11*, 591-596.
Meade, A. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational and Organizational Psychology, 77*, 531-552.
Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102,* 246-268.
Moskowitz, D. S. (2009). Coming full circle: Conceptualizing the study of interpersonal behaviour. *Canadian Psychology, 50,* 33-41.
Mulaik, S. A. (2005). Looking back on the indeterminacy problems in factor analysis. In A. Maydeu-Olivares, & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P.*

McDonald (pp. 153-172). Mahwah, NJ: Lawrence Erlbaum Associates.
Pace, V. L., & Brannick, M. T. (2010). How similar are personality scales of the "same" construct? A meta-analytic investigation. *Personality and Individual Differences, 49,* 669-676.
Pettersson, E., Turkheimer, E., Horn, E. E., & Menatti, A. R. (2011). The General Factor of Personality and evaluation. *European Journal of Personality, 26,* 292-302.
Revelle, W., & Wilt, J. (2009). *How important is the general factor of personality? A general critique.* Department of Psychology, Northwestern University. Evanston, IL. Retrieved January 15th, 2011 from http://personalityproject.org/revelle/syllabi/454/gfp.pdf.
Revelle, W. & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega and the glb: comments on Sijtsma. *Psychometrika. 74*, 145-154.
Riemann, R., & Kandler, C. (2010). Construct validation using multitrait-multimethod-twin data: The case of a general factor of personality. *European Journal of Personality, 24*, 258-277.
Rushton, J. P., Bons, T. A., Ando, J., Hur, Y-M., Irwing, P., Vernon, P. A., Petrides, K. V., & Barbaranelli, C. (2009). A general factor of personality from multitrait-multimethod data and cross-national twins. *Twin Research and Human Genetics, 12*, 356-365.
Rushton, J. P., Bons, T. A., & Hur, Y-M. (2008). The genetics and evolution of a general factor of personality. *Journal of Research in Personality, 42,* 1173-85.
Rushton J. P., & Irwing, P. (2011) The General Factor of Personality (pp. 132-161). In T. Chamorro-Premuzic, A. Furnham, & S. von Stumm (Eds.), *Handbook of Individual Differences*. London: Wiley-Blackwell.
Samuel, D. B., Simms, L. J., Clark, L. A., Livesley, W. J. & Widiger, T. A. (2010). An item response theory integration of normal and abnormal personality scales. *Personality Disorder: Theory, Research, and Treatment, 1*, 5-21.
Saville, P., & Wilson, E. (1991). The reliability and validity of normative and ipsative approaches in the measurement of

personality. *Journal of Occupational and Organizational Psychology, 64,* 219-238.
Tomarken, A. J., & Waller, N. G. (2003). Potential Problems With "Well Fitting" Models. *Journal of Abnormal Psychology, 112,* 578-598.
Vassend, O., & Skrondal, A. (2011). The NEO personality inventory revised (NEO-PI-R): Exploring the measurement structure and variants of the five-factor model. *Personality and Individual Differences, 50,* 1300-1304.
Veselka, L., Just, C., Jang, K. L., Johnson, A. M., & Vernon, P. A. (2012). The General Factor of Personality: A critical test. *Personality and Individual Differences, 52,* 261-264.
Veselka, L., Schermer, J. A., Petrides, K. V., Cherkas, L. F., Spector, T. D., & Vernon, P. A. (2009). A general factor of personality: Evidence from the HEXACO model and a measure of trait emotional intelligence. *Twin Research and Human Genetics, 12,* 420-424.
Veselka, L., Schermer, J. A., Petrides, K. V., & Vernon, P. A. (2009). Evidence for a heritable general factor of personality in two studies. *Twin Research and Human Genetics, 12,* 254-260.
Vinkhuyzen, A. A. E., van der Sluis, S., Maes, H. H. M., & Posthuma, D. (2012). Reconsidering the heritability of intelligence in adulthood: Taking assortative mating and cultural transmission into account. *Behavior Genetics, 42*, 187-198.
Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement. 9,* 1-26.
Widaman, K.F. (1993). Common Factor Analysis Versus Principal Components Analysis: Differential Bias in Representing Model Parameters? *Multivariate Behavioral Research, 28*, 263-311.
Wood D, & Roberts, B. W. (2006). Cross sectional and longitudinal tests of the personality and role identity structural model (PRISM). *Journal of Personality, 74,* 779-809.
Zawadzki, B., & Strelau, J. (2010). Structure of personality: The search for a general factor viewed from a temperament perspective. *Personality and Individual Differences, 49*, 77-82.

Chapter 6

Rushton: The Great Theoretician and His Contribution to Personality

Donald I. Templer
Alliant International University (Retired)
Fresno, California, USA
257 W. Los Altos
Fresno, CA 93704

Abstract

J. Philippe Rushton has been described as one of the greatest theorists in the field of personality and individual differences and social psychology. His life history theory (Differential K Theory) was said to bring unification to the scientific study of racial differences as Einstein's theory of relativity unified the field of physics. It was maintained that Rushton's General Factor of Personality (GFP) meshes with both general intelligence and his Differential K Theory. His work on the correlation of dark coloration and aggression is based on animal and human research. The present author presented a two-pillar theory of adaptive functioning that has a heavy concentration on Rushton conceptualization.

Introduction

There have been two books that were most influential in my career. When I was a clinical psychology graduate student in my twenties, I was unable to believe what American graduate students were supposed to believe. I could not believe that psychotherapy works very much. I could not believe that projective techniques had much validity. I could not feel a religious type fervor that people are capable of great change. Instead, I believed that our lives are very much determined by our biology. Psychology suddenly made sense to me when I read Hans Eysenck's (1957) *The Dynamics of Anxiety and Hysteria*. I did not realize at the time that my theoretical orientation was essentially that of the "English school" with roots in Darwin, Galton, Spearman, Pearson, Burt, Eysenck, and Jensen. The second influential book was Rushton's (1998) *Race, Evolution and Behavior: A Life History Perspective* read when I was in my 60's and a professor of psychology. I corresponded with him and met him at a conference of the International Society for Intelligence Research. He became my mentor in racial realism research.

I regard Rushton as one of the greatest theoreticians in the field of psychology. He is the great unifier. (1) He integrates psychology with the other social sciences, especially anthropology, history, and sociology. (2) He integrates psychology with the biological sciences, especially genetics, anatomy, and endocrinology. (3) He integrates psychology with animal behavior (and evolutionary theory.) (4) His life history theory (Differential K Theory) incorporates a vast array of behavioral, personality, biological, and social/ethnic differences. (5) His genetic similarity theory extends kin selection to the human case, integrates assortative mating and behavioral genetics, and explains ethnic nepotism. I will here repeat what I previously said, that in my opinion, "the unifying theory regarding racial differences of Rushton is the equivalent of Einstein's unification of physics by the theory of relativity" (Templer, 2004, p. 11).

Rushton, like the classical personality theorists of a century ago, is a very creative and prolific writer. That is where the

resemblance ends and the scientific superiority of Rushton dwarfs these early theorists. Rushton's concepts are more comprehensible, definable, and quantifiable. They are much more empirically verifiable and have been empirically verified. (Incidentally, I have never heard Rushton say that he is a personality theorist. Perhaps he does not want to be grouped with imprecise thinkers.) Rushton is extremely creative and writes extremely well. Not all psychologists who are creative and write well are outstanding researchers. Rushton is extremely proficient in statistics and research design. In fact, his classical article in *Psychological Bulletin* (Rushton, Brainerd, & Pressley, 1983) on aggregated data will endure independently of his work on intelligence and personality. In the Psych Lit system alone that article has well over 300 citations.

Rushton's Differential K Theory

The *r* vs. *K* differentiation of reproductive strategies and life histories of animals was introduced by MacArthur and Wilson (1967). The *r* strategists, in contrast to *K* strategists, have large litter size, shorter birth spacing, more offspring, less parental care, higher infant mortality, more rapid maturation, shorter life, higher reproductive effort, higher energy utilization, lower social organization, and lower altruism.

Rushton very capably described how Blacks and East Asians (Chinese, Japanese, and Koreans) throughout the world are at opposite ends of the *r-K* continuum with Blacks having more r characteristics and East Asians having more *K* characteristics. East Asians have larger brain size, higher intelligence based on different types of measurements, greater cultural achievements, slower motor development, slower skeletal development, slower dental development, older age of first intercourse, older age of first pregnancy, longer life span, lower aggressiveness, greater conscientiousness, lower dominance, less impulsivity, lower self-concept, lower sociability, higher marital stability, higher law abidingness, higher social organization, lower two-egg twinning,

lower hormonal levels, smaller genitals, smaller secondary sexual characteristics, less frequent intercourse, less permissive attitude toward sex, and lower rate of sexually transmitted diseases.

Einstein said that genius is the ability to see the obvious. Rushton's *r-K* conceptualization seems obvious after but not before one reads about it. The racial differences, however, are recognized by observant and open-minded people of average intelligence and education. They cannot help but notice that Blacks are greatly overrepresented in violent crime. They can't help but notice that Blacks talk louder and can be heard at a greater distance. They sing louder and play music more loudly. They are more muscular and have more of an athletic walk, are more athletically oriented and are athletically superior (Entine, 2000). Blacks talk more about sex and have more sexual awareness. On college campuses, East Asians are overrepresented in the library and underrepresented on the athletic fields.

East Asians are overrepresented in the more demanding disciplines—the physical sciences, mathematics, engineering, and pharmacy. The stereotype of difficulty of discipline is borne out in the research of Templer, Tomeo, Arikawa, and Robins (2002), who found a correlation of .76 between Verbal + Quantitative + Analytic Graduate Record Exam (GRE) and the ratio of Asians to Blacks in the discipline. The ratio was 18.16 in material engineering, 14.01 in physics, .45 in social work, and .11 in educational administration. California has three levels of state-supported educational institutions. The University of California schools are at the highest level, followed by the California State University campuses, followed by the community colleges. If I may exaggerate for the sake of emphasis, one almost feels like one is in an East Asian country at the University of California, Berkeley, with comparatively few Black students. This situation definitely does not prevail at the community colleges.

The race aspect of Rushton's life history theory may be the most controversial and most important with respect to social implications. It is, however, only one aspect of a very encompassing perspective. Life history theory involves sexual

behavior, number of offspring, parenting, health, life expectancy, impulse control, aggression, cooperation, and altruism. Intelligence and brain size are probably the most important determinants of this cluster of variables. In Rushton's theory, greater intelligence develops in colder climates because of the greater difficulty in obtaining and preserving food, making warm clothing, constructing shelters, and raising children to maturity.

Rushton stresses the role of testosterone in *r-K* and racial differences in aggressive and sexual behavior. Independently of Rushton's theory and findings, there is a plethora of research that shows higher testosterone level is associated with aggressive and risky behavior that includes criminal and especially violent criminal behavior, alcohol abuse, drug abuse, sexual promiscuity, interpersonal dominance, and physically risky behavior (Stanton et al., 2011).

Rushton's life history theory (Differential *K* Theory) is internally consistent. It makes sense. More importantly, there is good empirical support for Rushton's contention that greater intelligence develops in cold climates. There is also good empirical support for Rushton's contention that his important variables interrelated, as he said they should.

The research of Templer and Arikawa (2006) supports Rushton's contention that higher intelligence develops in colder climates. They used 129 countries with primarily indigenous populations (Asia, Africa, and Europe) and correlated culture fair IQ (Lynn & Vanhanen, 2002) with temperature. In addition to temperature, Templer and Arikawa used skin color provided by an anthropology book (Biasutti, 1967). IQ correlated -.92 with darker skin color, -.76 with winter highs, -.66 with winter lows, and .63 with per capita income. Previous research by Meisenberg (2004) reported a correlation of .89 between skin reflectance and IQ in a similar study. Templer (2010) found a correlation of .96 between skin color measure used by the Templer and Arikawa and the one by Meisenberg's skin reflectance. Because both measures were independently determined using different methodology, they both can be regarded as highly valid measures of skin color.

The research of Templer (2008) attempted to determine whether Rushton's *r-K* variables intercorrelate in the expected direction. Templer intercorrelated IQ, birth rate, infant mortality, HIV/AIDS, and life expectancy across 129 countries. All of the correlations were in the direction predicted on the basis of Differential *K* Theory. A *K* super factor accounted for 75% of the variance. The correlations were significantly higher with skin color, a biological variable, than with gross domestic product, a culturally influenced variable. The median of the 21 intercorrelations among the seven variables was .68.

Rushton's General Factor of Personality (GFP)

The GFP, like all of Rushton's ideas, is very much consistent with common sense. It is to personality as little *g* of Spearman and Jensen. And in the next section, we will see that the *Diagnostic and Statistical Manual of the American Psychiatric Association* has a 1-to-100 General Adaptive Functioning scale that ranges from being happy with excellent interpersonal relationships and contributing to society at one end to being completely dysfunctional or dangerous to self or others at the other. The roots of the GFP are in his statement that "one basic dimension—*K*—underlies much of the field of personality" (Rushton, 1985, p. 445). Rushton and Irwing (2009c, p. 356) succinctly described the person high in GFP as "altruistic, emotionally stable, agreeable, conscientious, extraverted, intellectually open, and mentally tough with high level of well-being, satisfaction with life, self-esteem, and emotionally intelligent," and they state that the person low in GFP is generally maladjusted and likely to have a personality disorder.

Considerable research has lent considerable credibility to the GFP. The study of Rushton and Irwing (2009a) is especially noteworthy because they used the Millon Clinical Multiaxial Inventory (MCMI), an instrument heavily oriented toward personality disorders. In fact, 13 of the 26 scales purport to assess personality disorders on the *Diagnostic and Statistical Manuel of the American Psychiatric Association*. In contrast, only one of the

10 regular MMPI scales was constructed to measure personality disorder. The MCMI yields a strong GFP.

The article of Rushton and Irwing (2009b) actually consisted of four different studies, one with the Big Five, one with the Gilford-Zimmerman Temperament Survey, one with the California Personality Inventory, and one with the Temperament and Character Inventory. All four are well established and validated personality inventories that are more geared for general populations than for psychiatric or other clinical populations. In all four studies, there was a solid GFP. Erdle, Irwing, Rushton, and Park (2010) found the Big Five and the Rosenberg Self-Esteem Scale yielded a substantial GFP.

The above rather cursory review of GFP research was to give the reader a quick grasp of the nature of the topic and findings. The review of Rushton and Irwing (2011) is much more extensive and detailed. They described the historical roots of the GFP and they pointed out that Darwin and Galton before them also viewed favorable personality characteristics as having evolved because they contributed to the success of the individual and society. The GFP correlations for identical twins are higher than for fraternal twins (Rushton & Irwing, 2008; Rushton & Irwing, et al., 2009c). Other research provides strong support for genetic influence (Figueredo, Vasquez, Brumbach, Schneider, 2004; Figueredo & Rushton, 2009). Finding material for GFP research is not difficult. "Tons" of relevant material has already been collected and is available for researchers. Rushton and associates have given us the big picture of the general structure. Nevertheless, "All of the cream has not been skimmed off the top." One area that needs more research is the personality deterioration seen in psychiatric disorders. The truth is bent when the general public is told that the mentally ill are no more violence prone than the average person. Also, antipsychotic drugs are less effective for the negative symptoms of schizophrenia than for the positive symptoms. Many parents are frustrated by their adult children who may no longer be hearing voices but have no motivation for work or school and spend all day watching television and smoking cigarettes. Many persons

with post-traumatic stress disorder are permanently incapacitated with respect to work and interpersonal relationships. The GFP has the capacity to bridge the gap between what personality psychologists write about and what clinical psychologists write about. Another area ripe for GFP research is economic success. One would think that people who are emotionally stable, likable, extraverted, open, and with good judgment would be more financially successful. In academia, eminence is measured by citations. In the military, status is determined by rank. In sports, athletes are assessed by number of points averaged. In business and industry, however, and as perceived by people in general, a person is judged by amount of money made and amount of financial worth. GFP would appear to have research applications for faculty in social and personality psychology and in business administration, economics, and organizational behavior. It is here suggested that research be carried out that relates GFP to IQ, income, net worth, and discrepancy between income and net worth. There are likely ethnic differences. East Asians are probably high in income and net worth and in ratio of net worth to income. In contrast, some American professional athletes make millions of dollars and yet have zero or negative net worth. An apparently relevant and useful instrument is the Money Attitude Scale (Templer &Yamauchi, 1982). It has five factors: Power-Prestige, Retention-time, Distrust, Quality, and Anxiety.

Psychologists should not ask how many personality factors there are. We tell our students that the results from factor analysis are a function of the nature of the population and the nature of the measuring instruments. We know that the number of factors for intellectual measures varies but that we can always count on obtaining g. In like fashion, we can always be certain of obtaining GFP with personality measures.

Two Pillars of General Adaptive Functioning

The following biological theoretical formulation is my own but is heavily influenced by the work of Rushton and associates. It is here

maintained that there are two biological pillars of general adaptive function. One is g of Spearman (1927) and Jensen (1998). The other is Rushton's GFP. If one assumes as suggested above that *r-K* is to group and races as GFP is to individuals, Rushton's Differential K Theory fits into my two-pillar theory.

Axis V of the D*iagnostic and Statistical Manual-IV (DSM-IV) of the American Psychiatric Association* is a 1-to-100 rating scale. "Superior functioning in a wide range of activities, life's problems never seem to get out of hand, is sought out by others because of his or her many positive qualities" (p. 34) merits a score from 91-to-100. "Persistent danger of severely hurting self or others (e.g., recurrent violence) or persistent inability to maintain minimum personal hygiene or serious suicidal act with clear expectation of death" (p. 34) merits a score of 0-10. The present author's formulation was intended to apply more to the big domain of adaptive function than to Axis V per se.

The two pillars are only slightly correlational with each other in high-functioning people. In fact, the California Psychological Inventory (CPI) Manual (Gough, 1992), presents 334 correlations between the CPI scales and cognitive/intellectual measures. Only 22 of these correlations are over .30. With below-average intelligence people, however, the two pillars are more highly correlated.

Herrnstein and Murray (1994) presented a plethora of evidence that people who are less bright function less well. Below-average IQ people do less well in school, are more likely to be unemployed, are more accident prone, are more likely to be physically disabled, commit more crime, and are more likely to be on welfare. In mental retardation, the inadequacy and dysfunction in behavior are even more prominent. One study found that only 13% of retarded children were free of psychiatric disorder and that 26% were psychotic (Gath & Gumley, 1986). Placement in institutions is often precipitated by maladaptive behavior such as hitting, kicking, biting, fecal and urinary incontinence, and public masturbation. Failure at job placement was often associated with aggressive or irritating or annoying behavior (Foss & Petersen,

1981). It should be noted that borderline and mild mental retardation are more likely to be of genetic etiology in contrast to moderate, severe, and profound retardation that are more likely to be caused by a medical condition or trauma early in life.

In most people and throughout their life history, the two pillars are determined primarily by genetics. Assault to the brain, however, can lower cognitive and personality functioning. These assaults to the brain could be short term, such as alcohol or drug intoxication. Brain disorders, however, are often permanent. Brain tumors can cause general intellectual decline, memory difficulty, personality deterioration, and psychosis. The early symptoms of Huntington's chorea are slovenly habits, irresponsibility, lower moral standards, and emotional lability. Next come choreic movements and general intellectual decline. In Alzheimer's disease, one progresses from forgetfulness to poor judgment, and then to dementia and incontinence and death. Both depression and psychoses are common (Wragg & Jeste, 1989). The symptoms of Pick's disease and Alzheimer's disease are similar, with the former having more personality deterioration and the latter more cognitive deterioration. Epileptics have a high suicide rate (Mathews & Barabas, 1981). A disproportionate number of persons in prisons are epileptic (Hermann & Whiteman, 1984). A quarter of epileptics with normal IQ have at least one psychiatric hospitalization (Stevens, Milstein, & Goldstein, 1972).

An interaction of genetics and cognitive functioning and personality/psychopathology is seen in schizophrenia. The premorbid IQ of persons with schizophrenia tends to be below average. This is probably because schizophrenic patients are more likely to have a history of prenatal and perinatal problem (McNeil & Kaij, 1973), head injuries (Wilcox & Nasrallah, 1987), abnormal EEG (Itil, Simeon, & Coffin, 1976), and thickness of the corpus callosum (Rosenthal & Bigelow, 1972). When schizophrenic persons are floridly psychotic, their cognitive functioning worsens. This is because of their looseness of associations, bizarre ideation, distracting internal stimuli such as voices, anxiety, depression, agitation, and fear of being hurt by other people. Antipsychotic

drugs not only lessen psychotic symptoms but bring about higher cognitive functioning (Cuesta, Jalon, Campos, & Peralta, 2009).

The structure under two pillars is supported by the combined strength of the two pillars. In psychology, as in civil engineering, weaknesses in one pillar can be compensated for by the other. Examples of higher intelligence predicting more favorable prognosis are seen in conduct disorder (Bassarath (2001), alcoholism (Gibbs & Flanagan, 1977), legal difficulty in drug abuse (Knapp, Templer, Cannon, & Dobson, 1991), mentally retarded sex offenders (Rice, Harris, Lang, & Chaplin, 2008), child molestation (Willis & Grace, 2008), and delinquent girls 10-14 years (Maskin, 1974).

How do life events fit into the two-pillar biological theory? Not even the most radical biological determinist can deny that such events as loss of a loved one, loss of a job, divorce, or disabling injury can have an effect on one's functioning. I view biology as determining one's life course and life events as producing deviations from the trajectory.

Dark Coloration and Aggression

The roots of Rushton relating aggressiveness to darker coloration are implicit in his Differential *K* Theory. *K* personality features and high intelligence evolve more in colder climates and persons in warm climates would be expected to have more r characteristics, which include aggressiveness, poor impulse control, low frustration tolerance, and showing less concern for the needs and rights of others. These people, as was discussed above, tend to be dark-complected. A more recent impetus of Rushton's work relating aggressiveness and dark coloration involves his incorporation of animal behavior. He was especially influenced by the article by Ducrest, Keller, and Roulin (2008) reporting darker color being associated with greater aggressiveness in 10 species of mammals, three species of birds, and several species of lizards.

Other animal research is consistent with the above-cited article by Ducrest et al. (2008). Chickens that lack black

pigmentation are more vulnerable to feather pecking from other chickens (Jensen, 2010). Blue tit birds with larger black patches were found to be more aggressive (Remy, Gregoire, Perret, & Doutrelant, 2010). Bokony, Garamszegi, Hirschenhauser, and Liker (2008) presented the findings of a number of studies showing correlations as high as .88 between melanization of plumage and testosterone level.

Rushton and Templer (2009), using 113 countries found that murder, rape, and serious assault were associated with darker skin color, lower IQ, higher birth rate, higher infant mortality, higher HIV/AIDS rate, lower life expectancy, and lower income. The correlations with violent crime were higher with skin color, a biological variable, than with income, a more culturally influenced variable. Templer and Rushton (2011) used 50 U.S. states and the same independent r-K variables as in the above Rushton and Templer (2009) study. All of the correlations were in the same direction. Darker skin correlated .84 with murder rate. Total violent crime correlated more highly with darker skin color ($r = .55$) than with income ($r = -.17$).

The study of Templer (in submission) is an extension of the work of Lynn (2010), who found that the Italian regions in northern Italy have higher mean IQ's than the southern regions. Lynn attributed his findings to the DNA of northern Italians resembling that of middle Europeans and that of southern Italians to North African and Middle Easterners. Templer determined biological correlates of IQ and found that cephalic index, hair color, eye color, multiple sclerosis rates, and schizophrenia rates were correlated with IQ. More germane to the present article is that attempted murder rate correlated .94 with percentage of black eyes, -.41 with percentage of blue eyes, -.45 with percentage of blond hair, and .48 with percentage of black eyes. Other researchers have found that White children with blue eyes are more behaviorally inhibited than those with brown eyes (Rosenberg & Kagan, 1987; Rosenberg & Kagan, 1989). European children with blue eyes and blond hair were found to be shyer than those with dark eyes and hair.

Further research relating eye and hair color in behavior and personality is recommended. It is here conjectured that a disproportionate number of White persons in contact sports such as boxing, wrestling, multiple martial arts fighting, football, rugby, and hockey are darker. Perhaps darker people go into more assertive occupations such as police officer, debt collector, bouncer, and prison guard. Of even more theoretical and practical importance would be comparing eye and hair color of violent and non-violent criminals. In general, and especially with adults, eye color would be a less methodological problematic variable than hair color. People grey with age and men bald with age. It is currently fashionable for men to shave their heads. Women, and men to a lesser extent, often dye their hair. It should be noted that in the Templer (in submission) study, hair and eye color were based on young Italian military recruits in the middle of the 19th century (Livi, 1896, 1905). Greying and balding and dying of hair were presumably not methodological complications.

John Calvin (1509-1564), the famous Protestant leader and theologian, believed in the predestination of salvation. Many Calvinists have maintained that wealth is an external manifestation of salvation. It is here contended that skin color is one external manifestation of intelligence and of prosocial vs. antisocial behavior. It is especially a good indicator on an international level and a long-term historical perspective. This generalization is less applicable on a group level. East Asians have higher IQ's and larger brains than Whites but are a little darker because of greater cloud cover in Europe than northeast Asia. Ashkenazi Jews, most of whom do not have blond hair and blue eyes, are the brightest people in the world, probably a function of primarily the brightest surviving centuries of gross maltreatment.

Conclusion

Previous psychologists obtained eminence by writing that early life events have a strong bearing on behavior and functioning of the individual. Rushton, although not disparaging such positions, went

further by contending that evolution of thousands of years influences present behavior and personality of individuals and of groups/races. Previous psychologists have obtained eminence by their work on intelligence and others by their work on personality. Rushton maintained that the same evolutionary process that determines intelligence determines personality and behavior. Lower GFP, more r life history characteristics, lower g, and lower two-pillar adaptive functioning tend to be associated with antisocial behavior, aggression, and crime. Skin color may be an external indicator of such undesirable behavior.

References

Bassarath, L. (2001). Conduct disorder: A biopsychosocial review. *Canadian Journal of Psychiatry, 46*(7), 609-616.
Biasutti, R. (1967). *Le Razze e I Popoli Della Terra.* Torino, Italy: Unione Tipografico-Editrice Torinese.
Bokony, V., Garamszegi, L. Z., Hirschenhauser, K., & Liker, A. (2008). Testosterone and melanin-based black plumage coloration: a comparative study. *Behavioral Ecology and Sociobiology, 62*, 1229-1238.
Coplan, R., Coleman, B., & Rubin, K. (1997). *Shyness and the little boy blue: Skin pigmentation, gender, and social wariness in preschoolers.* Hoboken, NJ: John Wiley & Sons, 37-44.
Cuesta, M. J., Jalon, E. D.., Campos, M. S., & Peralta, V. (2009). Cognitive effectiveness of olanzapine and risperidone in first-episode psychosis. *British Journal of Psychiatry, 194*, 439-445.
Ducrest, A., Keller, L., & Roulin, A. (2008). Pleitropy in the melanocortin system, coloration and behavioural syndromes. *Trends in Ecology and Evolution, 23*, 502-510.
Entine, J. (2000). *TABOO: Why Black athletes dominate sports and why we are afraid to talk about it.* New York: Public Affairs.
Erdle, S., Irwing, P., Rushton, J. P., & Park, J. (2010). The general factor of personality and its relation to self-esteem in 628,640 Internet respondents. *Personality and Individual Differences, 48*, 343-346.
Eysenck, H. J. (1957). *The dynamics of anxiety and hysteria: An experimental application of modern learning theory to psychiatry.* London: Routledge & Kegan Paul.
Farrington, D. P., Ttofi, M. M., & Coid, J. W. (2009). Development of adolescence-limited, late-onset, and persistent offenders from age 8 to age 48. *Aggressive Behavior, 35*(2), 150-163.
Figueredo, A. J., & Rushton, J. P. (2009). Evidence for shared genetic dominance between the general factor of personality,

mental and physical health, and life history traits. *Twin Research and Human Genetics, 12*(6), 555-563.
Figueredo, A. J., Vasquez, G., Brumbach, B. H., & Schneider, S. M. R. (2004). The inheritability of life history strategy: The *K*-factor, covitality, and personality. *Social Biology, 51*, 121-143.
Foss, G., & Peterson, S. L. (1981). Social-interpersonal skills relevant to job tenure for mentally retarded adults. *Mental Retardation, 19*(3), 103-106.
Gath, A., & Gumley, D. (1986). Behaviour problems in retarded children with special reference to Down's syndrome. *British Journal of Psychiatry,* 149, 156-161.
Gibbs, L., & Flanagan, J. (1977). Prognostic indicators of alcoholism treatment outcome. *International Journal of the Addictions, 12*, 1097-1141.
Gough, H. G. (1987). *California Psychological Inventory Administrator's Guide*. Palo Alto, CA: Consulting Psychologists Press, Inc.
Hermann, B. P., & Whiteman, S. (1984). Behavioral and personality correlates of epilepsy: A review, methodological critique, and conceptual model. *Psychological Bulletin, 95*(3), 451-497.
Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and class structure in American life.* New York: Free Press.
Holland, T. R., & Holt, N. (1975). Prisoner intellectual and personality correlates of offense severity and recidivism probability. *Journal of Clinical Psychology, 31*(4), 667-672.
Itil, T. M., Simeon, J., & Coffin, C. (1976). Qualitative and quantitative EEG in psychotic children. *Diseases of the Nervous System, 5*, 147-252.
Jensen, A. R. (1998). *The g factor.* Westport, CT: Praeger.
Jensen, P. (2010). Domestication, selection, behavior and welfare of animals—genetic mechanisms for rapid responses. *Animal Welfare, 19*(S), 7-9.
Kagan, R. (1989). Physical and physiological correlates of behavioral inhibition. *Developmental Psychology, 22*(8), 753-761.

Knapp, J. E., Templer, D. I., Cannon, W. G., Dobson, S. (1991). Variables associated with success in an adolescent drug treatment program. *Adolescence, 26*(102), 305-307.
Livi, R. (1896, 1905). Antropometria Militare. Risulati ottenuti dallo spoglio dei fogli sanitari dei military delle classi 1859-1863 eseguito dall'Ispettorato di sanita militare per ordine del Ministero della Guerra, 1, Data antropologici ed etnologici, Atlante della geografia antropologica d'Italia. Roma: Giornale medico del Regio Esercito.
Lynn, R. (2008). *The global Bell Curve. Race, IQ, and inequality worldwide*. Augusta, GA: Washington Summit Publishers.
Lynn, R. (2010). In Italy, North-South differences in IQ predict differences in income, education, infant mortality, stature, and literacy. *Intelligence, 38*, 93-100.
Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations*. Westport, CT: Praeger.
Maskin, M. B. (1974). A comparison of graduate and recidivist WISC IQ scores in a delinquency treatment program for girls. *Journal of Clinical Psychology, 30*(3), 319-210.
Mathews, M. S., & Barabas, G. (1981). Suicide in epilepsy: A review of the literature. *Psychosomatics, 22*, 515-524.
MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, NJ: Princeton University Press.
McNeil, T. F., & Kaij, L. (1973). Obstetric complications and physical size of offspring of schizophrenic, schizophrenic-like, and control mothers. *British Journal of Psychiatry, 123*(574), 341-348.
Meisenberg, G. (2004). Talent, character and the dimensions of national culture. *Mankind Quarterly, 45*, 123-168.
Remy, A., Gregoire, A., Perret, P., & Doutrelant, C. (2010). Mediating male-male interactions: the role of the UV blue crest coloration in blue tits. *Behavioral Ecology and Sociobiology, 64*, 1839-1847.
Rice, M. E., Harris, G. T., Lang, C., & Chaplin, T. C. (2008). The sexual preferences and recidivism of sex offenders with mental retardation. *Sexual Abuse, 20*(4), 409-425.

Rosenberg, A. A., & Kagan, J. (1987). Iris pigmentation and behavioral inhibition. *Developmental Psychobiology, 20*(4), 377-392.
Rosenberg, A. A., & Kagan, J. (1989). Physical and physiological correlates of behavioral inhibition. *Developmental Psychobiology, 22*(8), 753-770.
Rosenthal, R., & Bigelow, L. B. (1972). Quantitative brain measurements in chronic schizophrenia. *British Journal of Psychiatry, 121*(562), 259-264.
Rushton, J. P. (1985). Differential K theory: The sociobiology of individual and group differences. *Personality and Individual Differences, 6*, 441-452.
Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective*. New York: Brunswick, NJ: Transaction.
Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity. The principle of aggregation. *Psychological Bulletin, 94*(1), 18-38.
Rushton, J. P., & Irwing, P. (2009a). A general factor of personality in the Millon Clinical Multiaxial Inventory-III, the Dimensional Assessment of Personality Pathology, and the Personality Assessment Inventory. *Journal of Research in Personality, 43*, 1091-1095.
Rushton, J. P., & Irwing, P. (2009b). A general factor of personality in 16 sets of the Big Five, the Guilford-Zimmerman Temperament Survey, the California Psychological Inventory, and the Temperament and Character Inventory. *Personality and Individual Differences, 47*, 558-564.
Rushton, J. P., & Irwing, P. (2009c). A General Factor of Personality (GFP) from the Multidimensional Personality Questionnaire. *Personality and Individual Differences, 47*, 571-576.
Rushton, J. P., & Irwing, P. (2011). The General Factor of Personality: Normal and abnormal. In T. Chamorro-Premuzic, S. von Strumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences*. Hoboken, NJ: Wiley-Blackwell.

Rushton, J. P., & Templer, D. I. (2009). National differences in intelligence, crime, income, and skin color. *Intelligence, 37*, 341-346.
Spearman, C. (1927). *The Abilities of Man.* New York: Macmillan.
Stanton, S. J., O'Dhaniel, O. A., Mullette-Gillman, McLaurin, R. E., Kuhn, C. M., LaBar, K. S., Platt, M. L., & Huettel, S. A. (2011). Low- and high-testosterone individuals exhibit decreased aversion to economic risk. *Psychological Science, 22*, 447-453.
Stevens, J. R., Milstein, V., & Goldstein, S. (1972). Psychometric test performance in relation to the psychopathology of epilepsy. *Archives of General Psychiatry, 26*, 532-538.
Templer, D. I. (2004). The Bell Curve: An assessment after ten years. *The Occidental Quarterly*, 4(3), 7-20.
Templer, D. I. (2008). Correlational and factor analytic support for Rushton's differential *K* life history theory. *Personality and Individual Differences, 45*, 440-444.
Templer, D. I. (2010). IQ and skin color: The old world reexamined and the new world. *Mankind Quarterly, 51*, 3-25.
Templer, D. I., & Arikawa, H. (2006). Temperature, skin color, per capita income, and IQ: An international perspective. *Intelligence, 34*, 121-139.
Templer, D. I., & Rushton, J. P. (2011). IQ, skin color, crime, HIV/AIDS, and income in 50 U.S. states. *Intelligence, 39*(20), 437-442.
Templer, D. I., Tomeo, M. E., Arikawa, H., & Williams, R. (2003). Asian-Black differences in aptitude and difficulty of chosen academic discipline. *Personality and Individual Differences, 35*, 237-241.
Wilcox, J. A., & Nasrallah, H. A. (1987). Childhood head trauma and psychosis. *Psychiatry Research, 21*(4), 303-306.
Willis, G. M., & Grace, R. C. (2008). The quality of community reintegration planning for child molesters: Effects on sexual recidivism. *Sexual Abuse, 20*(2), 218-240.

Wragg, R. E., & Jeste, D. V. (1989). Overview of depression and psychosis in Alzheimer's disease. *American Journal of Psychiatry, 146*(5), 577-587.

Yamauchi, K., & Templer, D. I. (1982). The development of a Money Attitude Scale. *Journal of Personality Assessment, 46*(5), 522-528.

Chapter 7

J. P. Rushton's Contribution to the Study of Altruism

Yoon-Mi Hur
Mokpo National University, South Korea

Abstract

The main purpose of this essay is to describe Rushton's work on altruism over the past 40 years. During his academic career, he changed his theoretical approach from social learning theory to sociobiology. This essay includes four sections. The first section gives an overview of Rushton's work on altruism. The second section reviews Rushton's early work on altruism based on social learning theory. His laboratory and naturalistic experiments led him to conclude that altruism could be increased by exposure to models exemplifying the behavior, and that, once engaged in, it could be durable over months and generalizable across age groups and situations. The third section will discuss his more recent work on the genetic foundation of altruism with social biology theory and empirical evidence. Rushton conducted twin studies of prosocial behaviors using both western and Asian samples across different age groups. He concluded that about 50% of the variance in prosocial behaviors is heritable. The last section will describe altruism as part of the general factor of personality (GFP), which occupies the apex of the hierarchy of personality.

Introduction

I am honored to be a contributor on this Special Issue for J. P. Rushton. I have been fortunate to work with Rushton on papers about genetic influences on prosocial behaviors in South Korean children and other research on personality .His contribution to knowledge has become well-known over the past four decades. Like other eminent psychologists, Rushton has boldly challenged conventional views and the minds of mediocrity. As a result, he naturally had to endure strong oppositions from critics in the field of psychology and public opinion. But, nevertheless, his work has also been acknowledged with admiration as well. He is one of the most provocative and controversial figures in the recent history of psychology.

Altruism has been a central focus for his research during his entire career. His study of altruism began with his doctoral dissertation at the University of London (completed in 1973). During the 1970s and 1980s, many psychologists focused on negative traits such as aggression and delinquency. However, Rushton maintained the importance of also studying positive aspects of human nature. With conviction that human beings are helpful, cooperative, empathic, loving, and considerate, he asserted that altruism is a universal value in all human societies and forms the basic tenet of most of the world's great religious, social reformist, and revolutionist movements (Rushton, 1982). He defined altruism as "social behavior carried out to achieve positive outcomes for another rather than for self." In this formulation, egoism is the opposite end of the continuum from altruism. In defining altruism, Rushton emphasized the behavioral component, although he did not preclude psychological motivations activating the behavior such as "empathy" and "moral judgment," which he thought were primary (Rushton, 1980, 1982).

Rushton's work on altruism has spanned a considerable range of methodologies. Early in his career (1970-1980), he performed many laboratory and naturalistic field experiments to study altruism from the social learning perspective. He also

contributed to the settlement of the dispute on "consistency vs. specificity" (Mischel 1968) by applying the principle of aggregation to measure altruistic and other behaviors. Later, he became deeply fascinated with the theory of evolution and social biology and shifted to study altruism from the evolutionary and sociobiological perspectives. In this paper I first review Rushton's early work on altruism performed mostly on the basis of social learning theory, and then discuss his more recent work on the genetic foundation of altruism using the empirical evidence he provided. Finally, I will talk about altruism in the context of the research on the general factor of personality (GFP). I will use the terms 'prosocial behaviors' and 'altruism' interchangeably throughout this paper.

Early work: A Social learning perspective on altruism

Rushton (1976, 1980) contended that the extent to which a person engages in altruistic behavior and the motivations underlying the behavior are the results of the person's social learning experiences. In other words, a person is honest, generous, helpful, and compassionate to the degree to which he has learned to be so. Rushton particularly emphasized that families, educational systems, and mass media are major agencies of socialization in society. Rushton (1982) criticized problems of under-socialization in society, however. He argued that the family was becoming an ineffective socializer of children due to the break-up of the traditional extended family, and that TV was too often socializing children in an antisocial rather than prosocial direction. Rushton provided constructive solutions to the problem of under-socialization as well. Rushton (1980, 1982) suggested that we should increase the frequency of parents-child interactions, introduce stronger regulatory guidelines to stop the pervasiveness of antisocial socialization on TV, and develop courses of psychology and sociology as well as more intensive and disciplined programs of prosocial education in schools.

During the 1970s, Rushton conducted many laboratory studies and naturalistic field experiments to demonstrate the effects of observational learning, reinforcement learning, and learning from direct verbal procedures such as instructions and preaching.. For example, Emler and Rushton (1974) found that 7- to 13-year-olds gave more valuable tokens to a charity after listening to stories concerned with distributive justice. The number they gave was significantly related to their level of moral judgment measured as responses to the stories. In a related experiment, Rushton (1975) demonstrated that children exposed to a generous model, who donated tokens to the charity, donated significantly more than children exposed to a selfish mode. The modeling effects lasted over several weeks. Models who preached selfishness produced less giving behaviors in children than did models who preached generosity or neutral message. Other studies showed that positive reinforcement following model-induced altruistic behaviors led to increased generosity in children while punishment led to a decrement (Grusec,Kuczynski, Rushton, & Simutis; 1978). Rushton and Campbell (1977) examined the modeling of prosocial behaviors in adults on blood donating. In that field experiment, he showed that observing a model not only affected an immediate test of volunteering to donate blood but also affected whether blood was actually donated 6 weeks later in a naturalistic setting.

From a review of experimental studies conducted by himself and others, Rushton (1976, 1980) made three big conclusions about modeling effects. First, exposing subjects to altruistically behaving models impacted the amount and direction of those subjects' subsequent altruistic behaviors. Secondly, these model-induced altruistic behaviors lasted over weeks. Finally, modeling effects were generalized across different age groups and situations.

During the 1960s and 1970s, a debate on "consistency vs. specificity" in human behaviors predominated in the area of personality assessment. Mischel (1968) had fervently denied the existence of personality traits, arguing that the coefficient of behaviors across different situations is only about +.30. While

many personality psychologists had a pessimistic view of finding a stable trait of altruism, Rushton (1980, 1981) and Rushton, Brainerd, and Presley (1983) reviewed the personality literature and demonstrated that consistent patterns of individual differences are found if composite measures are used . Rushton (1980, 1981) also found significant correlations among the measures of altruism by self-report, peer-rating, and whether students completed an organ-donor card attached to their drivers' license. These observations led Rushton to conclude that a broad-based "altruistic personality' does exist and that altruism can be validly and reliably measured by self-report as well as by objective methods.

The genetic foundation of altruism

In the late 1970s and early 1980s, Rushton underwent a paradigm shift from social learning theory to sociobiology. He began to think that altruistic behaviors like helping, protecting, and nurturing could occur in the absence of previous learning experiences. He sought the origin of altruism from evolutionary history, suggesting that altruism is an inborn-part of our species-typical repertoire. Sociobiology explains altruistic behaviors by the principle of 'inclusive fitness' and 'kin selection' (Hamilton, 1964). If the net result of an altruistic act helps more of that individual's genes to survive (even in other people such as kin) and to be transmitted to future generations, then the act is adaptive even if it results in the death of the individual. Furthermore, in order to replicate their own genes more effectively, organisms act to benefit relatives depending on how closely related they are. Paradoxically, altruism has a 'selfish' purpose.

Altruism can be found in nonhuman animals as well as in humans. In nonhuman animals, altruism includes parental care, warning calls, cooperative defense, rescue behavior, food sharing, and self-sacrifice. Rushton (1980), gave the good example of altruism found in a honeybee: the poisonous sting of the bee is an adaptation against hive robbers. The recurved barbs facing backward from the sharp tip cause the whole sting to be wrenched

out of the bee's body, along with some of the bee's vital internal organs. These barbs have been described as instruments of altruistic self-sacrifice. Although the individual dies, the bee's genes survive. Parenting is another altruistic behavior that has evolved through differential reproductive success. In a population of social animals, those that protect their offspring will successfully raise more progeny than those who do not. Any genes associated with parenting behaviors will thereby out reproduce the selfish alternatives. Rushton (1991) hypothesized that empathy in humans is a disposition that might have evolved from parental care. Numerous twin studies have documented that genes influence individual differences in empathy (Bouchard & McGue, 2003; Hur, 2007; Rushton et al., 1986).

Rushton (1989) proposed genetic similarity theory to extend the idea of kin selection and explain altruism toward nonrelatives among humans, for example, spouse selection and friendship formation. In a study of adult twins and spouses, Rushton and Bons (2005) found that for the responses to personality and attitude questions the correlations were .53 for monozygotic twin pairs, .32 for dizygotic twins, .32 for spouse pairs, and .20 for pairs of best friends. Moreover, they found that partner similarity was more pronounced on the more heritable items than it was on the less heritable items. Taken together, these results suggest that people are genetically inclined to choose as social partners those who resemble themselves at the genetic level. Rushton (1991) also explained 'ethnic nepotism' from the view of genetic similarity theory. People tend to help members of their own race or ethnic group more than they do members of other races or foreigners. Rushton contends this is because individuals within an ethnic group are genetically more similar to each other than they are to individuals from different ethnic groups.

Rushton and his colleagues have conducted several twin studies and found that altruism and prosocial behavior are heritable. For example, Rushton, Fulker, Neale, Nias, and Eysenck (1986) found that 50% of the variance of altruism in 573 pairs of adult twins was associated with genetic effects, with the remainder being

due to environmental effects. The authors also found that the important environmental factors influencing individual difference in altruism were not those shared by siblings such as social class, parental values, or child rearing styles but were those unique to each member of a twin pair. Rushton (1981, 2004) included "social responsibility" as part of altruism. He administered a 22-item Social Responsibility Questionnaire to a subsample of adult twins who participated in his earlier twin study (Rushton et al., 1986) and found that 42% of the reliable variance was due to genes, confirming strong genetic influences. In contrast to the earlier work (Rushton et al., 1986), however, Rushton (2004) noted that 23% of the reliable variance of Social Responsibility was due to the twins' common family environment. These findings were consistent with a study by Krueger et al. (2001) where the authors found a significant proportion of common family environmental variance in prosocial behaviors.

Because prosocial behaviors are in part rooted in genes, it is reasonable to expect some connections between sex hormones and prosocial behaviors. In an attempt to detect underlying physiological mechanisms for genetic influences on prosocial behaviors, Rushton and colleagues (Harris, Hampson, Rushton & Jackson, 1996) found that among university students testosterone was significantly negatively related with prosocial behaviors, while it was significantly positively related to aggression in both males and females. Given genetic influences on the level of testosterone found in other studies (Hoekstra, Bartels, & Boomsma, 2006), these results point out a direction to study whether common genes operate in the correlations between the testosterone level and human social behaviors.

Several studies have found that children as young as 15 months spontaneously and repeatedly give things to each other and that babies who hear other babies' cries begin crying themselves, suggesting that prosocial behaviors and empathy emerge very early in childhood (e.g., Zahn-Waxler, Schiro, Robinson, Emde, & Schmitz, 2001). Rushton and I (Hur & Rushton, 2007) examined mothers' ratings of prosocial behavior in 514 pairs of 2- to 9-year-

old South Korean twins. Monozygotic and dizygotic twin correlations showed a tendency of increasing genetic effects and decreasing family environmental effects with age. Variance component analyses revealed that 55% of the total variance of prosocial behaviors was associated with genetic influences, with the remainder due to environmental effects unique to each member of a twin pair and measurement error. Common family environmental influences were near zero. Overall, these results were similar to those reported by studies of prosocial behaviors based on western samples (e.g., Knafo & Plomin, 2006; Rushton et al., 1986). Important implications of these findings are two-fold: First, genetic influences on prosocial behavior emerge very early in childhood. Secondly, genetic factors in prosocial behaviors found in western samples may be generalized to Asian populations.

More recently, using 1110 pairs of South Korean adolescent and young adult twins [mean age = 18.0y (SD = 3.3y)], Rushton and I (Hur, Jeong, Schermer, & Rushton, 2011) explored genetic and environmental influences on miserliness, a trait opposite to generosity. Freudian theory, which is still cited in some psychology textbooks, suggests that miserliness originates from a fixation at the anal stage in childhood, that is, about two years of age (Kalat, 2011). Thus, if a child receives too much pressure or punishment from parents during toilet training, he will experience anxiety over bowel movement and subsequently take pleasure in withholding feces, which may lead to the later development of miserliness. However, our analyses demonstrated that family environmental factors were negligible for the development of miserliness, whereas genetic influences were about 28%. Non-genetic variance was due to individual-specific environmental influences and measurement error.

Altruism and the General Factor of Personality

Rushton's work on altruism and personality has been integrated into a very broad hypothesis known as the General Factor of Personality (GFP). According to this view, both narrow and broad

personality traits (miserliness, generosity, altruism) can be organized hierarchically with the GFP at the apex of both "normal" personality traits (i.e., non-clinical) and clinical personality disorders, in the same way in which g, the general factor of mental ability, occupies the apex in the organization of cognitive abilities (Rushton, Bons, & Hur, 2008; Rushton & Irwing, 2011). High scores on the GFP indicate what is meant by someone having a "good" personality; low scores indicate what is meant by a "difficult" personality, in other words someone who is hard to get along with. Individuals high on the GFP are altruistic, agreeable, relaxed, conscientious, sociable, and open, with high levels of well-being and self-esteem. Because the GFP defines clear positive and negative poles, it provides potential for understanding the socially "advantaged" (those with high levels of emotional intelligence) as well as the socially "challenged" (those more likely to suffer a personality disorder).

The GFP has been found across diverse samples, procedures, and personality inventories. For example, one study found the GFP was independent of method variance using a multitrait–multimethod analysis of self-, teacher-, and parent-ratings of 391 13- to 14-year-olds using the Big Five Questionnaire—Children (Rushton et al., 2009). Several cross-national twin studies have found 50% of the variance on the GFP attributable to genetic influence and 50% to non-shared environmental influence. A South Korean twin study showed the GFP had emerged by 2- to 3-years of age (Rushton et al., 2008).

The GFP can be viewed as a dimension of social effectiveness. The explanation Rushton et al. (2008) suggested for the GFP is that it arose through evolutionary selection for socially desirable traits that facilitate performance across a wide range of contexts. This followed a proposal by Charles Darwin that natural selection acted directionally, to endow people with more cooperative and less contentious personalities than their archaic ancestors or nearest living relatives, the chimpanzees. Rushton et al. (2008) conjectured that individuals high on the GFP left more progeny, since people prefer as mates, fellow workers, and leaders

those who are altruistic, conscientious, and emotionally stable. People able to cooperate in groups were also more likely to win competitions and wars.

References

Bouchard, T. J. Jr., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology, 54*, 4–45.

Emler, N. P., & Rushton, J. P. (1974). Cognitive-developmental factors in children's generosity. *British Journal of Social and Clinical Psychology, 13*, 277-281.

Grusec, J. E., Kuczynski, L., Rushton, J. P., & Simutis, Z. M. (1978). Modeling, direct instruction, and attributions: Effects on altruism. *Developmental Psychology, 14*, 51-57.

Hamilton, W. D. (1964). The genetical evolution of social behavior: I and II. *Journal of Theoretical Biology, 7*, 1-52.

Hoekstra, R. A., Bartels, M., & Boomsma, D. I. (2006). Heritability of testosterone levels in 12-year-old twins and its relations to pubertal development. *Twin Research and Human Genetics, 4*, 558-565.

Krueger RF, Hicks BM, McGue M. (2001). Altruism and antisocial behavior: independent tendencies, unique personality correlates, distinct etiologies. *Psychological Science, 12*(5):397-402.

Harris, J. A., Rushton, J. P., Hampson, E., & Jackson, D. N. (1996). Salivary testosterone and self-report aggressive and pro-social personality characteristics in men and women. *Aggressive Behavior, 22*, 321-331.

Hur, Y. M., Jeong, H-K., Schermer. J. A., & Rushton, J. P. (2011). Miserliness is heritable. *Personality and Individual Differences, 51*, 1052-1055.

Hur, Y.-M., & Rushton, J. P. (2007). Genetic and environmental contributions to prosocial behavior in 2- to 9-year-old South Korean twins. *Biology Letters, 3*, 664-666.

Kalat, W. J. (2011). *Introduction to psychology*, 9th Ed. Belmont, CA: Wadworth.

Knafo, A. & Plomin, R. (2006). Prosocial behaviour from early to middle childhood: Genetic and environmental influences on stability and change. *Developmental Psychology, 42*, 771-786.
Mischel, W. (1968). *Personality and assessment*. Mahwah, NJ: Lawrence Erlbaum
Rushton, J. P. (1973). *Social learning and cognitive development: Alternative approaches to an understanding of generosity in 7 to 11 year olds*. Unpublished doctoral dissertation, University of London, England.
Rushton, J. P. (1975). Generosity in children: Immediate and long term effects of modelling, preaching, and moral judgment. *Journal of Personality and Social Psychology, 3*, 459-466.
Rushton, J. P. (1976). Socialization and the altruistic behavior of children. *Psychological Bulletin, 83*, 898-913.
Rushton, J. P., & Campbell, A. C. (1977). Modeling, vicarious reinforcement, and extraversion on blood donating in adults: Immediate and long term effects. *European Journal of Social Psychology, 7*, 297-306.
Rushton, J. P. (1980). *Altruism, socialization, and society*. Englewood Cliffs, NJ: Prentice Hall.
Rushton, J. P. (1981). The altruistic personality. In J. P. Rushton & R. M. Sorrentino (Eds.). *Altruism and helping behaviour: Social, personality, and developmental perspectives* (pp. 251-266). Hillsdale, NJ: Erlbaum.
Rushton, J. P. (1982). Altruism and society: A social-learning perspective. *Ethics, 92*, 425-446.
Rushton, J. P. (1989). Genetic similarity, human altruism, and group selection (with commentaries and author's response). *Behavioral and Brain Sciences, 12*, 503-559.
Rushton, J. P. (1991). Is altruism innate? *Psychological Inquiry, 2*, 141-143.
Rushton, J. P. (2009). Inclusive fitness in human relationships. *Biological Journal of the Linnean Society, 96*, 8-12.
Rushton, J. P., Bons, T. A., Ando, J., Hur, Y-M., Irwing, P., Vernon, P. A., Petrides, K. V., & Barbaranelli, C. (2009). A general

factor of personality from multitrait-multimethod data and cross-national twins. *Twin Research and Human Genetics, 12*, 356-365.

Rushton, J. P., & Bons, T. A. (2005). Mate choice and friendship in twins: Evidence for genetic similarity. *Psychological Science, 16*, 555-559.

Rushton, J. P., Bons, T. A., & Hur, Y-M. (2008). The genetics and evolution of a general factor of personality. *Journal of Research in Personality, 42,* 1173-85. (Corrigendum, 2009, *Journal of Research in Personality, 43,* 532.)

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94,* 18-38.

Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences, 2,* 293-302.

Rushton, J. P., Fulker, D. W., Neale, M. C., Nias, D. K. B., & Eysenck, H. J. (1986). Altruism and aggression: The heritability of individual differences. *Journal of Personality and Social Psychology, 50*, 1192-1198

Rushton, J. P., & Irwing, P. (2011). The general factor of personality: Normal and abnormal. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell Handbook of Individual Differences* (pp. 132-161). London: Wiley-Blackwell.

Zahn-Waxler, C., Schiro, K., Robinson, J. L., Emde, R. N., & Schmitz, S. (2001). Empathy and prosocial patterns in young MZ and DZ twins: Development and genetic and environmental influences. In R. N. Emde & J. K. Hewitt (Eds.), *Infancy to early childhood: Genetic and environmental influences on developmental change* (pp. 141-162). New York: Oxford University Press.

Chapter 8

The Measurement of Human Life History Strategy

Aurelio José Figueredo [a,⇑], Tomás Cabeza de Baca [b], Michael Anthony Woodley [a]

[a] *Department of Psychology, University of Arizona, Tucson, AZ, USA*

[b] *Family Studies and Human Development, University of Arizona, Tucson, AZ, USA*

Abstract

This retrospective essay appraises J. Philippe Rushton's application of life history to understanding the covariation among human traits in light of subsequent developments in the measurement and latent structure of Human Life History, covitality, and personality. We conclude that Rushton should be recognized for having initiated a theoretically and empirically highly fertile paradigm in human individual differences research.

1. Introduction

Differential Psychology (DP) is the study of individual differences, Behavioral Genetics (BG) is the study of their heritability, and Evolutionary Psychology (EP) is the study of how their frequencies change over generations. Although they have since drifted apart, Sir Francis Galton envisaged them as complementary facets of a unified, coherent field of study (Rushton, 1990a).

One key difference between DP and EP involves the dynamics underlying traits. DP presupposes the existence of individual differences, proceeding from there to explore their implications for a person's life, work, and social relationships. BG also takes the existence of heritable variations for granted, proceeding from there to trace their origins in our genetic material. First and foremost, EP views traits as strategic, meaning differentially selected in phylogenies based on their consequences for survival and reproduction. Natural Selection represents the differential survival and reproduction of variant individuals resulting from what Darwin (1859) called their differential abilities to combat the "hostile forces of nature", meaning threats of morbidity and mortality from factors such as cold, drought, predators, parasites, and even prey. Sexual Selection instead entails the differential reproduction of variant individuals resulting from what Darwin (1871) androcentrically called their differential abilities to "charm the females" and "to conquer other males in battle", meaning to obtain and retain partners for sexual reproduction. Finally, Social Selection denotes differential success in social competition or cooperation with members of one's own species, regardless of the sex. This specific term was coined later (Nesse, 2007; West-Eberhard, 1979) because Darwin (1871) had described the process in detail but failed to label it uniquely.

Selection by Consequences provides EP with the perspective that traits may be adaptations to specific environmental contingencies (adaptive problems). Similarly, the genetic effects studied by BG are not taken as a given, but instead as shaped by selective pressures over evolutionary time.

Although few evolutionary psychologists will express these ideas bluntly, most of them are likely to agree with this representation, perhaps cautiously adding that not all traits can be presumed to represent evolved adaptations, but might instead be side-effects of adaptations or even selectively neutral. What remains controversial is at what level one should investigate these hypothesized adaptive functions. One currently popular view is that they should each be studied piecemeal, because the human mind is modular and each trait functions as a semi-autonomous entity that is separately shaped by its own unique selective pressures (Barkow, Cosmides, & Tooby, 1992). An alternative view is that one is well advised to examine the selective consequences of how traits, such as personality, risk taking, and health, are combined into trait clusters that may be differentially selected based on how well they do or do not work together to serve their multiple adaptive functions (Rushton, 1985a).

For the past three decades, a pioneer of this second point of view has been J. Philippe Rushton, who was among the first to see how Life History (LH) theory (MacArthur & Wilson, 1967; Pianka, 1970) could be productively applied to make sense of the otherwise seemingly incomprehensible clustering of traits that we observe. LH theory describes the ways organisms allocate re-sources among different components of fitness, given the environmental constraints under which they survive and reproduce. LH strategies range from those based on maximizing reproductive outcomes as a buffer against environmental unpredictability, to those based on maximizing longevity and parenting, so as to enhance the quality and competitiveness of organisms living in stable environments (Ellis, Figueredo, Brumbach, & Schlomer, 2009). LH strategies have traditionally been characterized as r-selected and K-selected, representing opposite poles of a continuum. The former comprises species exhibiting very fast life histories (e.g., rabbits), with r denoting the maximum population reproductive rate of that species. The latter encompasses species exhibiting very slow life histories (e.g., elephants), which saturate their environment to carrying capacity (K). For example, a high mating effort strategy represents

a fast life history owing to the tendency for enhanced reproduction to be associated with more rapid maturation and to come at the expense of longevity, whereas the high somatic effort and parental effort strategy represents a slow life history, owing to its theoretical and empirical associations with slower ontogenic development and enhanced longevity (Figueredo, Vásquez, Brumbach, & Schneider, 2004; Figueredo et al., 2005).

Rushton (1985a) made a series of innovative predictions concerning the relationship between individual differences and LH strategy, a substantial number of which have been empirically validated by subsequent work. Rushton's original (1985a) paper initiated an alternative research tradition in LH theory (See: Rushton, 1985b, 1985c, 1987a, 1987b, 1988a, 1988b, 1990a, 1990b, 1991, 1992, 2000, 2004; Bogaert & Rushton, 1989; Figueredo & Rushton, 2009; Rushton & Ankney, 1993; Rushton & Bogaert, 1988; Rushton & Templer, 2009; Rushton & Whitney, 2002; Templer, 2008; Templer & Rushton, 2011). This bold idea proposed that hierarchically organized and heritable individual differences variables such as personality and intelligence share a common source of variance stemming from life history, which acts to co-ordinate tradeoffs among traits. He proposed that heritable variations in LH strategy are not only at the root of human individual differences but are also at the root of ethnic/racial and socioeconomic status differences along the same parameters.

Perhaps the most significant of these claims is that LH strategy can be measured as an individual differences variable in its own right. Bogaert and Rushton (1989) were the first to explore this possibility with their life history questionnaire, among whose 18 constituent measures a latent LH factor was identified. Bogaert and Rushton included items directly measuring ecological factors such as density, dispersion and competition, in addition to attitudinal and physiological factors. This approach integrated with earlier density-dependent conceptualizations of LH evolution (Pianka, 1970), as these are measures of individual differences in ecological preferences.

Whereas many of us originally trained in the biological sciences were quite conversant in LH theory before encountering the work of Rushton, we simply did not envision all the implications he derived for human differential psychology. For example Rushton (1985a) derived the following prediction: "An exciting if open ended possibility is that one basic dimension — K — underlies much of the field of personality" (p. 445). Consistent with this theoretical expectation, a General Factor of Personality (GFP) has now been extracted from over 24 different personality inventories (Rushton & Irwing, 2011) and is predicted to exist in others (Woodley & Bell, 2011). Rushton (1985a) also predicted that health outcomes (such as disease resistance and longevity) might share a common variance with LH strategy, owing to the organismic quality versus quantity tradeoff implicit in LH theory. Consistent with this prediction is the finding of a latent Covitality factor in biomedical and subjective well-being measures (Weiss, King, & Enns, 2002). Contemporary measures of LH strategy, not based on density-dependent selection, have been found to correlate so strongly with both the GFP and this Covitality Factor that a higher-order factor has been extracted explaining the preponderance of the covariance among the three lower-order factors (e.g., Figueredo, Vásquez, Brumbach, & Schneider, 2006; Figueredo, Vásquez, Brumbach, & Schneider, 2007; Figueredo et al., 2004; Figueredo et al., 2005). The existence of this higher-order LH factor validates some very important predictions made by Rushton (1985a), in that it has been found to exhibit a modestly high heritability (h^2 .65; Figueredo et al., 2004; Figueredo et al., 2005), which corroborates Rushton's prediction that human LH is a biologically prepared disposition rather than simply a product of socialization, as previous researchers had proposed. The modestly high heritability of this higher-order LH factor further accords with Rushton's utilization of the standard assumption of BG, namely that very few traits are purely heritable and that all individual difference traits are also modestly environmentally plastic. Indeed, consistent with this position, much recent research indicates that humans are able to adaptively calibrate their life histories in

response to perceptions of environmental stability during key phases of development (Del Giudice & Belsky, 2011; Ellis et al., 2009).

Prior to the work of Rushton, explaining why there exists so much observable covariation among socially problematic behaviors (e.g., timing of intercourse, alcohol/drug usage, etc.) had perplexed social scientists. Most explanations offered were proximate, meaning immediately causal or developmental mechanisms for these life outcomes, ignoring any ultimate explanations regarding the possible adaptive significance of this covariation (Cabeza de Baca, Figueredo, & Ellis, in press). The prevailing proximate explanation was that social privilege, meaning the institutional advantages a family possessed (e.g., due to interethnic social dominance, higher income, higher social class, etc.) could account for behavioral variation both between and within groups (Figueredo et al., 2007).

In the 1960s and 1970s, however, the sociobiology movement urged social scientists to incorporate ultimate or evolutionary explanations in the study of human behavior (e.g., Dawkins, 1976; Hamilton, 1964; Trivers, 1972; Williams, 1966; Wilson, 1975), proposing that the social sciences could be unified into a broader biological sciences framework conducive to the generation of novel distal-level hypotheses (Cabeza de Baca et al., in press; King & Cabeza de Baca, 2011). Unfortunately, this synthesis was met with resistance from the social sciences, citing melioristic reasons, such as the perpetuation of inequity and the justification for deprived conditions among disadvantaged groups (Charlesworth, 1992; King & Cabeza de Baca, 2011). Despite resistances, social science started taking notice of the new evolutionary framework that sociobiology evoked. By the late 1970s and 1980s, nuanced approaches of sociobiology began to emerge. Specifically, researchers began to examine physiological, personality and other individual differences contextually, noting that, beyond human universals, adaptations were context-specific. Weinrich (1977) examined the variation between racial and socioeconomic classes differences concerning sexual behavior and

patterns of pair bonding – citing that resource unpredictability, conceptualized as family income, could differentially impact sexual decision-making. By 1983, Reynolds and Tanner (1983) began incorporating context via religious practices and societal culture, systematically reviewing differences in conception, adolescence, and marriage, among others. Also, in 1983, a workshop discussing the merits of LH theory on human research was convened by representatives from anthropology, psychology, and biology (Weigel & Blurton-Jones, 1983).

Aside from these early precursors, we maintain that it was the work of Rushton that gave direction to nearly all that followed. Unfortunately, his work on the psychometrics of LH strategy was overshadowed by the controversy surrounding his prediction of race differences in LH strategy as adaptations to ancestral ecological conditions in the different regions of the world.

2. The psychometrics of Human Life History

Contemporary LH theory suggests that there might be a number of factors, independent of density-dependent selection, that give rise to differences in life history speed. Contemporary measures of human LH strategy incorporate these newer perspectives while elaborating the nomological net first established by Rushton (1985a) for the specific implications of LH strategy in our species.

2.1. Cognitive and behavioral indicators of slow life history strategy

Using an undergraduate student sample, Figueredo et al. (2005) estimated a slow LH factor that loaded positively and significantly on Attachment to and Investment from their Biological Father and Adult Romantic Partner Attachment, and loaded negatively and significantly on Attachment to and Investment from any "Other" Father Figure, the Mating Effort Scale, Machiavellianism Scale, and Risk Taking Questionnaire. Using an ethnically homogeneous sample of Mexican, non-student adults with children of their own, Tal, Hill, Figueredo, Frías-Armenta, and Corral-Verdugo (2006)

estimated a slow LH factor that loaded positively and significantly on Long-Term Planning, Father's Parental Investment, Mother's Parental Investment, Parental Investment Towards Children, Social Contact and Support from Family and from Friends, Responsibility and Support for Kin and for Non-Kin.

The 20 cognitive and behavioral scales identified from within the National Survey for Midlife Development in the US (MIDUS; Brim et al., 2000) data by Figueredo et al. (2007) as theoretically predicted convergent indicators of a latent variable (K) representing slow LH strategy, measured the following psychosocial traits: Agency, Advice Seeking, Foresight/Anticipation, Insight Into Past, Primary Control/Persistence, Flexible/Positive Reappraisal, Self Directedness/Planning, Financial Status, Health Control, Mother and Father Relationship Quality, Marital Relationship Quality, Children Relationship Quality, Family and Friends Social Support, Altruism Towards Kin and Non-Kin, Close Relationship Quality, Communitarian Beliefs, and Religiosity. All of these scales had common factor loadings from the slow LH factor ranging from .30 to .60. In the MIDUS twins (Figueredo et al., 2004), all of these psychosocial traits had genetic factor loadings from the slow LH factor ranging from .50 to 1.00, except for Agency and Health Control, which had no significant genetic covariance.

Using a composite college student sample, Sefcek and Figueredo (2010) subsequently constructed a slow LH factor loading positively and significantly on the Mini-K (Figueredo et al., 2006) (A 20-item short form of the 199-item, Arizona Life History Battery or ALHB; Figueredo, 2007), Mother and Father Parental Investment, Family Social Support, Friends Social Support, Romantic Partner Attachment, and General Altruism. Gladden, Figueredo, and Jacobs (2008) obtained similar findings with an independent student sample, estimating a slow LH factor loading positively Mother and Father Parental Investment, Family Social Support, Friends Social Support, Romantic Partner Attachment, and General Altruism, and Religiosity. This particular set of cognitive and behavioral indicators constituted the ALHB

(Figueredo, 2007), which has been used in a variety of subsequent studies.

With respect to criterion-related validity, a variety of investigators have found that using either this entire set or some restricted subset of cognitive and behavioral indicators, one can: (1) positively and significantly predict Executive Functions, Trait Emotional Intelligence, Mate Value Inventory, Mate Value Scale, Rosenberg Self Esteem Scale, Self-Adjective Checklist, Social Economic Exchange Scale, Collective Self Esteem Scale, Positive Assortative Mating, Female Physical Height, Long-Term Mating Sociosexual Orientation, Moral Intuitions, In-Group Loyalty, as well as Secure Attachment, Supportive Communication, and Long-Term Satisfaction in Romantic Relationships, both cross-sectionally and longitudinally; and (2) negatively and significantly predict the Mating Effort Scale, Short-Term Mating Sociosexual Orientation, Escalated Mate Retention Tactics, Affective and Punitive Responses to Sexual or Emotional Infidelity, Intimate Partner Violence, Interpersonal Aggression, Female Intrasexual Competitiveness, Disordered Eating Behavior, Negative Ethnocentrism, Negative Androcentrism, Levenson's Primary Psychopathy Scale, Levenson's Secondary Psychopathy Scale, Machiavellianism Scale, Buss-Perry Aggression Questionnaire, Proactive-Reactive Aggression Questionnaire, and General Social Deviance (e.g., Buunk, Pollet, Klavina, Figueredo, & Dijkstra, 2009; Figueredo, Andrzejczak, Jones, SmithCastro, & Montero-Rojas, 2011; Figueredo, Gladden, & Beck, 2011; Figueredo, Gladden, & Hohman, 2011; Figueredo & Wolf, 2009; Gladden, Sisco, & Figueredo, 2008; Gladden, Figueredo, & Snyder, 2010; Jones, Figueredo, Dickey, & Jacobs, 2007; Kirsner, Figueredo, & Jacobs, 2009; Olderbak & Figueredo, 2009; Olderbak & Figueredo, 2010; Salmon, Figueredo, & Woodburn, 2009; Wenner, Figueredo, Rushton, & Jacobs, 2007).

2.2. Slow life history and the general factor of personality

As previously noted, the General Factor of Personality (GFP) predicted by Rushton has now been extracted from over 24

different personality inventories (Rushton & Irwing, 2011), is robust to meta-analysis (van der Linden, Scholte, Cillessen, te Nijenhuis, & Segers, 2010), and exhibits a heritability of around .50 (Rushton, Bons, & Hur, 2008). Like the g-factor in intelligence research, the GFP is believed to exist at the apex of a hierarchy of personality traits, with Digman (1997) and DeYoung et al.'s (2002) Big Two traits (Stability or Alpha; and Plasticity or Beta) occupying the next level and Costa and McCrae's (1992) Big Five (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism) occupying the next level, and so on (Rushton & Irwing, 2011). The GFP also exhibits modest criterion related validity (van der Linden, te Nijenhuis, & Bakker, 2010; van der Linden, Bakker, & Serlie, 2011). 2.3. Slow life history and covitality

2.3. Slow life history and covitality

Because slow LH Strategists, by definition, allocate more bioenergetic and material resources to somatic effort, meaning growth and self-maintenance, we expect them to generally manifest better health and life expectancy. These individuals expend more effort on their personal condition and also receive more parental effort and nepotistic effort from family members (who share genes biasing development towards slow LH strategy), as well as more mutualistic and reciprocal altruism from non-relatives, including romantic partners, with whom they are more likely to develop long-term cooperative relationships.

This theoretical prediction has been tested and confirmed in a number of studies. Using data from MIDUS, Figueredo et al. (2004) related a latent variable representing slow LH strategy (K), composed of 20 convergent indicator scales, to a general factor representing a state of overall physical and mental health (Covitality). The latter construct loaded positively on Subjective Well-Being, Positive Affect, and General Health, and loaded negatively on Negative Affect and Medical Symptoms.

The MIDUS data were collected from telephone interviews and two follow-up mail surveys given to a nationally representative

sample in 1995–1996, sampled only 25–74 year-old English speakers in the US, and also contained a genetically informative sample of MZ and DZ twins.

In the subsample of singletons, the phenotypic correlation between the slow LH and Covitality factors was .50. In the subsample of twins, the phenotypic correlation between the slow LH and Covitality factors was .54, and the genetic correlation between the slow LH and Covitality factors was .69. All these correlations were highly statistically significant. Figueredo and Rushton (2009) followed up with more sophisticated biometric structural equations modeling and found that a higher-order LH factor termed Super-K loaded on slow LH, Personality and on this same Covitality factor (.63). This common factor variance was decomposed into 17% additive genetic variance, 8% attributable to shared environmental influences among twins raised together, and 36% attributable to unmeasured environmental influences that were not shared among them. The other indicators of the Super-K factor manifested no significant variance component attributable to shared environmental influences. Unlike these other LH indicators, however, the Covitality factor showed no significant non-additive genetic variance component attributable to "dominance" effects, meaning multiplicative gene-gene interactions.

Using a composite student sample, Sefcek and Figueredo (2010) subsequently constructed another Covitality factor, which loaded .31 on the slow LH factor. This Covitality factor loaded positively and significantly on the General health scale, the MOS SF-36 Short Form Health Survey, the Subjective Well-Being Scale, but negatively and significantly on the Medical Symptoms Scale, the Hopkins Anxiety Index, the Hopkins Depression Index, and the Beck Depression Inventory.

3. Conclusions

The application of life history theory to DP has developed into a sophisticated and nuanced analysis that has gained widespread utilization in diverse research ranging from pubertal timing (Ellis,

2004), to parenting (Cabeza de Baca et al., in press; Sotomayor-Peterson, Cabeza de Baca, Figueredo, & Smith-Castro, in press), to religiosity (Gladden, Welch, Figueredo, & Jacobs, 2009). Widespread application of LH theory would not have been possible without researchers within their respective fields reformatting their existing measures and conceptualizing their hypotheses to incorporate LH theory. Nonetheless, we expect more progress will be made toward the measurement of LH.

It is evident that Rushton's application of LH theory to understanding individual and group differences constitutes a significant and novel contribution to explaining the observed covariation among human behaviors. Rushton made many bold predictions and most have been supported to varying degrees by the preponderance of research that followed in his scientific wake. Rushton's work must further be praised for its willingness to engage with controversy. While not accorded the fullest general recognition that his work deserves in this current day and age, ultimately posterity will be the judge of the value of his work.

References

Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* New York: Oxford University Press.

Bogaert, A. F., & Rushton, J. P. (1989). Sexuality, delinquency, and r/K reproductive strategies: Data from a Canadian university sample. *Personality and Individual Differences, 10*, 1071–1077.

Brim, O. G., Baltes, P. B., Bumpass, L. L., Cleary, P. D., Featherman, D. L., Hazzard, W. R., Kessler, R, C., Lachman, M. E., Markus, H. R., Marmot, M. G., Rossi, A. S., Ryff, C. D. & Shweder, R. A. (2000). National Survey of Midlife Development in the United States (MIDUS), 1995–1996 [Computer file]. ICPSR version. DataStat and Harvard Medical School, Dept. of Health Care Policy [producers], 1996. Ann Arbor: Inter-University Consortium for Political and Social Research [distributor].

Buunk, A. P., Pollet, T. V., Klavina, L., Figueredo, A. J., & Dijkstra, P. (2009). Height among women is curvilinearly related to life history strategy. *Evolutionary Psychology, 7(4)*, 545–559.

Cabeza de Baca, T., Figueredo, A.J., & Ellis, B.J. (in press). An evolutionary analysis of variation in parental effort: Determinants and Assessment. *Parenting: Science and Practice.*

Charlesworth, W. R. (1992). Darwin and developmental psychology: Past and Present. *Developmental psychology, 28*, 5–16.

Costa, P. T., Jr., & McCrae, R. R. (1992). *The revised NEO personality inventory (NEOPIR) and NEO five-factor inventory (NEO-FFI) professional manual.* Odessa, FL.: Psychological Assessment Resources.

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life* (1st ed.). London: John Murray.

Darwin, C. (1871). *The descent of man, and selection in relation to sex* (1st ed.). London: John Murray.

Dawkins, R. (1976). *The selfish gene.* Oxford, UK: Oxford University Press.
Del Giudice, M., & Belsky, J. (2011). The development of life history strategies: Toward a multi-stage theory. In D. M. Buss & P. H. Hawley (Eds.), *The Evolution of Personality and Individual Differences* (pp. 154–176). New York: Oxford University Press.
DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences, 33*, 533–552.
Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246–1256.
Ellis, B. J. (2004). Timing of pubertal maturation in girls: An integrated life history approach. *Psychological Bulletin, 130*, 920–958.
Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Human Nature, 20*, 204–298.
Figueredo, A. J., Andrzejczak, D. J., Jones, D. N., Smith-Castro, V., & Montero-Rojas, E. (2011). Reproductive strategy and ethnic conflict: Slow life history as a protective factor against negative ethnocentrism in two contemporary societies. *Journal of Social, Evolutionary, and Cultural Psychology, 5*, 14–31.
Figueredo, A. J., Gladden, P. R., & Beck, C. J. A. (2011). Intimate partner violence and life history strategy. In A. Goetz & T. Shackelford (Eds.), *The Oxford Handbook of Sexual Conflict In Humans*, Chapter 5 (pp. 72–99). New York, NY: Oxford University Press.
Figueredo, A. J., Gladden, P. R., & Hohman, Z. (2011). The evolutionary psychology of criminal behavior. In S. C. Roberts (Ed.), *Applied Evolutionary Psychology*, Chapter 13 (pp. 201–221). New York, NY: Oxford University Press.
Figueredo, A. J., & Rushton, J. P. (2009). Evidence for shared genetic dominance between the general factor of personality, mental and physical health, and life history traits. *Twin Research and Human Genetics, 12*, 555–563.

Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. R. (2004). The heritability of life history strategy: The K-factor, covitality, and personality. *Social Biology, 51,* 121–143.
Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. R. (2007). The Kfactor, covitality, and personality: A psychometric test of life history theory. *Human Nature, 18*, 47–73.
Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M. R., Sefcek, J. A., Tal, I. R., et al. (2006). Consillience and life history theory: From genes to brain to reproductive strategy. *Developmental Review, 26*, 243–275.
Figueredo, A. J., Vásquez, G., Brumbach, B. H., Sefcek, J. A., Kirsner, B. R., & Jacobs, W. J. (2005). The K-Factor: Individual differences in life history strategy. *Personality and Individual Differences, 39*, 1349–1360.
Figueredo, A. J., & Wolf, P. S. A. (2009). Assortative pairing and life history strategy: A cross-cultural study. *Human Nature, 20*, 317–330.
Gladden, P. R., Figueredo, A. J., & Jacobs, W. J. (2008). Life history strategy, psychopathic attitudes, personality, and general intelligence. *Personality and Individual Differences, 46*, 270–275.
Gladden, P. R., Figueredo, A. J., & Snyder, B. (2010). Life history strategy and evaluative self-assessment. *Personality and Individual Differences, 48*, 731–735.
Gladden, P. R., Sisco, M., & Figueredo, A. J. (2008). Sexual coercion and life history strategy. *Evolution and Human Behavior, 29*, 319–326.
Gladden, P. R., Welch, J., Figueredo, A. J., & Jacobs, W. J. (2009). Moral intuitions and religiosity as spuriously correlated life history traits. *Journal of Evolutionary Psychology, 7(2),* 167–184.
Hamilton, W. D. (1964). The genetical evolution of social behavior I and II. *Journal of Theoretical Biology, 7*, 1–52.
Jones, D. N., Figueredo, A. J., Dickey, E. D., & Jacobs, W. J. (2007). Relations among individual differences in reproductive strategies, sexual attractiveness, affective and punitive intentions, and imagined sexual or emotional infidelity. *Evolutionary Psychology, 5*, 367–390.

King, A.C., & Cabeza de Baca, T. (2011). The stagnancy of family studies in modern academia: Resistances toward the integration of evolutionary theory. *Evolution: Education and Outreach, 4*, 64–74.
Kirsner, B. R., Figueredo, A. J., & Jacobs, W. J. (2009). Structural relations among negative affect, mate value, and mating effort. *Evolutionary Psychology, 7*, 374–397.
MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton. N.J: Princeton University Press.
Nesse, R. M. (2007). Runaway social selection for displays of partner value and altruism. *Biological Theory, 2*, 143–155.
Olderbak, S. G., & Figueredo, A. J. (2009). Predicting romantic relationship satisfaction from life history strategy. *Personality and Individual Differences, 46*, 604–610.
Olderbak, S. G., & Figueredo, A. J. (2010). Life history strategy as a longitudinal predictor of relationship satisfaction and dissolution. *Personality and Individual Differences, 49*, 234–239.
Pianka, E.R. (1970). On r- and K-selection. *American Naturalist, 104*, 592–596.
Reynolds, V., & Tanner, R. E. S. (1983). *The Biology of Religion*. New York: Longman.
Rushton, J. P. (1985a). Differential K theory: The sociobiology of individual and group differences. *Personality and Individual Differences, 6*, 441–452.
Rushton, J. P. (1985b). Differential K theory and group differences in intelligence. *Behavioral and Brain Sciences, 8*, 239–240.
Rushton, J. P. (1985c). Differential K theory and race differences in E and N. *Personality and Individual Differences, 6*, 769–770.
Rushton, J. P. (1987a). An evolutionary theory of health, longevity, and personality: Sociobiology and r/K reproductive strategies. *Psychological Reports, 60*, 539–549.
Rushton, J. P. (1987b). Toward a theory of human multiple birthing: Sociobiology and r/K reproductive strategies. *Acta Geneticae Medicae et Gemellologiae, 36*, 289–296.
Rushton, J. P. (1988a). Do r/K reproductive strategies apply to human differences? *Human Ethology Newsletter, 5*, 4–6.

Rushton, J. P. (1988b). Do r/K reproductive strategies apply to human differences? [Review of P. A. Racey & A. S. I. Loudon (Eds.): *Reproductive energetics in mammals*. Oxford, Oxford University Press, 1987.]. Social Biology, 35, 337–340.
Rushton, J. P. (1990a). Sir Francis Galton, epigenetic rules, genetic similarity theory, and Human Life History analysis. *Journal of Personality, 58*, 117–140.
Rushton, J. P. (1990b). Race differences, r/K theory, and a reply to Flynn. *The Psychologist: Bulletin of the British Psychological Society, 5*, 195–198.
Rushton, J. P. (1991). Do r-K strategies underlie human race differences? A reply to Weizmann et al. *Canadian Psychology, 32*, 29–42.
Rushton, J. P. (1992). Life-history comparisons between Orientals and Whites at a Canadian university. *Personality and Individual Differences, 13,* 439–442.
Rushton, J. P. (2000). *Race, evolution and behavior: A life history perspective* (3rd ed.). Port Huron: Charles Darwin Research Institute.
Rushton, J. P. (2004). Placing intelligence into an evolutionary framework, or how g fits into the r-K matrix of life history traits including longevity. *Intelligence, 32*, 321–328.
Rushton, J.P., & Irwing, P. (2011). The General Factor of Personality: Normal and abnormal. In T. Chamorro-Premuzic, S. von Strumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp. 134–163). Blackwell Publishing.
Rushton, J. P., & Ankney, C. D. (1993). The evolutionary selection of human races: A response to Miller. *Personality and Individual Differences, 15*, 677–680.
Rushton, J. P., & Bogaert, A. F. (1988). Race versus social class differences in sexual behavior: A follow up test of the r/K dimension. *Journal of Research in Personality, 22*, 259–272.
Rushton, J. P., Bons, T. A., & Hur, Y.-M. (2008). The genetics and evolution of the general factor of personality. *Journal of Research in Personality, 42*, 1173–1185.

Rushton, J. P., & Templer, D. I. (2009). National differences in intelligence, crime, income, and skin color. *Intelligence, 37*, 341–346.
Rushton, J. P., & Whitney, G. (2002). Cross-national variation in violent crime rates: Race, r-K theory, and income. *Population and Environment, 23*, 501–511.
Salmon, C., Figueredo, A. J., & Woodburn, L. (2009). Life history strategy and disordered eating behavior. *Evolutionary Psychology, 7*, 585–600.
Sefcek, J. A., & Figueredo, A. J. (2010). A life-history model of human fitness indicators. *Biodemography and Social Biology, 56*, 41–66.
Sotomayor-Peterson, M., Cabeza de Baca, T., Figueredo, A.J., & Smith-Castro, V. (in press). Coparenting, total parental effort, and life history strategy: A crosscultural comparison. *Journal of Cross-Cultural Psychology.*
Tal, I. R., Hill, D., Figueredo, A. J., Frías-Armenta, M., & Corral-Verdugo, V. (2006). An evolutionary approach to explaining water conservation. *Medio Ambiente y Comportamiento Humano, 7*, 7–27.
Templer, D. I. (2008). Correlational and factor analytic support for Rushton's differential K life history theory. *Personality and Individual Differences, 45*, 440–444.
Templer, D. I., & Rushton, J. P. (2011). IQ, skin color, crime, HIV/AIDS, and income in 50 U.S. states. *Intelligence, 39*, 437–442.
Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual Selection and the Descent of Man* (pp. 136–179). London: Heinemann.
van der Linden, D., Bakker, A. B., & Serlie, A. W. (2011). The General Factor of Personality in selection and assessment samples. *Personality and Individual Differences, 51*, 641–645.
van der Linden, D., Scholte, R. H. J., Cillessen, A. H. N., te Nijenhuis, J., & Segers, E. (2010). The general factor of personality and classroom ratings of likeability and popularity. *Journal of Research in Personality, 44*, 669–672.

van der Linden, D., te Nijenhuis, J., & Bakker, A. (2010). The General Factor of Personality: A meta-analysis of Big Five intercorrelations and a criterion related validity study. *Journal of Research in Personality, 44,* 315–327.
Weigel, R. W., & Blurton-Jones, N. G. (1983). Workshop report: Evolutionary lifehistory analysis of human behavior. *Behavioral Ecology and Sociobiology, 2,* 233–235.
Weinrich, J. D. (1977). Human socio-biology: Pair-bonding and resource predictability (effects of social class and race). *Behavioral Ecology and Sociobiology, 2,* 91–118.
Weiss, A., King, J. E., & Enns, R. M. (2002). Subjective well-being is heritable and genetically correlated with dominance in chimpanzees. *Journal of Personality and Social Psychology, 83,* 1141–1149.
Wenner, C., Figueredo, A. J., Rushton, J. P., & Jacobs, W. J. (2007). Executive functions, general intelligence, life history, psychopathic attitudes, and deviant behavior. Amsterdam, The Netherlands: Paper Presented at the 8th Annual Meeting of the International Society for Intelligence Research.
West-Eberhard, M. (1979). Sexual selection, social competition, and evolution. *Proceedings of the American Philosophical Society, 123,* 222–234.
Williams, G. C. (1966). *Adaptations and natural selection: A critique of some current evolutionary thought.* Princeton, NJ: Oxford University Press.
Wilson, E. O. (1975). *Sociobiology: The new synthesis.* Belknap: Cambridge, MA.
Woodley, M. A., & Bell, E. (2011). Is collective intelligence (mostly) the General Factor of Personality? *Intelligence, 39,* 79–81.

Web reference
Figueredo, A.J. (2007). The Arizona Life History Battery [Electronic Version]. <http:// www.u.arizona.edu/'ajf/alhb.html.>.

Chapter 9

J. P. Rushton's Theory of Ethnic Nepotism

Frank Salter [a,*], Henry Harpending [b]
[a] *Department of Government and International Relations, University of Sydney, Merewether Bld., Room 443, Sydney, Australia*
[b] *Department of Anthropology, University of Utah, Salt Lake City, USA*

Abstract

Unreciprocated aid among co-ethnics and the emotional intensity of ethnic conflict have long been explanatory challenges to evolutionary science. J. P. Rushton's theory of assortative ethnic affiliation–altruism, mating and friendship directed towards fellow ethnics–derives from his more general theory of genetic similarity (GST). GST proposes that humans give preferential treatment to others in whom they detect genetic resemblance and that such behaviour enhances genetic fitness. The theory coincides with W. D. Hamilton's theory of inclusive fitness as applied to relations between populations. GST helps explain core features of ethnicity, including its basis in putative kinship and correlation with gene frequencies. Ethnic nepotism due to similarity is a weak social force compared to social identity. However its pervasiveness makes it a potential driver of evolutionary and social change, a potential borne out by sociological studies of the impact of ethnic diversity on social cohesion and public altruism. Genomics confirms the theory for interactions within populations with sufficient genetic diversity,

such as ethnically mixed societies. GST applied to ethnicity is promising for further research in evolutionary social science because it unifies evolutionary and behavioral mechanisms in a single theory.

1. Theoretical background

The evolutionary approach to ethnicity has given priority to understanding affiliation outside the family and clan. Why is it that people so often prefer to marry and befriend fellow ethnics and defend their ethnic groups? These tendencies are present in all populations, indicating an evolutionary origin. But it has not been clear how such altruism could be evolutionarily stable. For ethnic solidarity to have evolved there must have been a substantial fitness payoff for sacrificing individual fitness for groups or populations that had ethnic characteristics.

An ethnic group is a named population whose members share a belief in common descent, have a shared history, a distinctive shared culture, a shared attachment to a homeland, and some degree of solidarity (Smith, 1986, pp.22-30). The core elements of this definition, including putative kinship, derive from the sociologist Max Weber (1946/1922, p.173). Genetic assay data show that ethnic kinship is real (e.g. Cavalli-Sforza, Menozzi & Piazza, 1994, p.75, Table; Harpending, 2002).

An early evolution theory of ethnic solidarity was provided by Eibl-Eibesfeldt (1972/1970; 1982), who argued that ethnic ties are based on family bonds. Behavioral adaptations for bonding with close kin came to be applied to whole populations, first bands, then tribes and nations. Affiliation to the hunter-gatherer band was subjected to group selection. The first part of this argument was reinterpreted sociobiologically by van den Berghe (1981), using Hamilton's theory of inclusive fitness. In his classic 1964 paper, Hamilton showed that genes coding for altruism would spread if altruism was directed towards genealogical kin, and if the resulting boost to their fitness exceeded the fitness cost to the altruist. This

condition for the evolution of altruism, known as Hamilton's Rule, was widely accepted by zoologists by 1980.

Like Eibl-Eibesfeldt, van den Berghe argued that because ethnicity is putative kinship, shared ethnic identity should release some of the same altruistic motivation found within families. He adopted Hamilton's concept of kin-recognition markers, which included language, territory, religion, and phenotypic similarity based on shared culture (language, clothing or scarification) and physical appearance. These recognition markers, which overlap those proposed by Weber, had evolved as releasers of nepotism because in the evolutionary environment they signaled kinship, though with different degrees of reliability (see also Shaw and Wong, 1989). Van den Berghe's theory appeared when sociobiological studies of animal altruism and kinship were appearing. This was the background against which Rushton and colleagues proposed their own theory.

2. Rushton's contribution

In 1984 Rushton and co-authors R. J. H. Russell and H. G. Wells proposed an evolutionary theory of ethnic altruism based on Genetic Similarity Theory (GST). GST seeks to explain a number of prosocial behaviors including ethnic affiliation, by generalizing Hamilton's theory such that genetic similarity alone elicits affiliative behavior without knowledge of genealogical kinship.

> "[W]e propose genetic similarity detection as a mechanism by which organisms are attracted and repelled by each other. We hypothesize that genetically similar others ("strangers," as well as "kin") have a tendency to seek each other out and provide mutually supportive environments, while genetically dissimilar others have a tendency to form natural antipathies and provide mutually hostile environments" (Rushton et al., 1984, pp.179-80).

The theory included evolutionary causality by asserting that responding differentially to genetic similarity increases an organism's fitness, defined as increased genetic representation in the population. It allowed for altruism among kin as well as among similar strangers. The theory held that assortment is stronger for the more heritable characteristics because they are more reliable indicators of genetic similarity, based on sociometrical data as well as a quantitative model by W. D. Hamilton (1996/1971).

The resulting theory of ethnic affiliation was more behaviorally detailed than Eibl-Eibesfeldt's and more general than van den Berghe's. Whereas social science research into ethnicity had been published in journals of sociology and anthropology, the new approach shifted the locus to psychology, ethology, and behavior genetics. Attention began to shift to genetics and selection theory.

Rushton continued to develop GST over the following years. In 1985 the team of Rushton, Russell and Wells again presented GST and argued that assortative mating for personality influenced the evolution of variation in human personality types. They combined GST with reciprocal altruism theory to predict that genetic similarity facilitates reciprocity by reducing the condition of complete reciprocity (Rushton et al., 1985, p.80).

The major statement of the relation between ethnic nepotism and genetic similarity was a paper by Rushton in *Behavioral and Brain Sciences* (Rushton, 1989a). One addition was data indicating that similarity among spouses and friends is strongest in the most heritable characteristics within a trait category. For example, married pairs are more alike in general intelligence *g* than they are in particular cognitive abilities such as vocabulary or arithmetic, and *g* is generally the more heritable cognitive ability. The same difference applies to the overall assemblage of traits on which spouses are similar, including anthropometrics (Rushton, 1989a, p.534; Russell & Rushton, 1985).

In the *BBS* article Rushton speculated about how similarity could give rise to ethnocentric ideology. He reviewed the literature

on extended phenotypic effects of genes, focusing on the idea that epigenetic developmental rules can incline people towards constructing and learning ideologies that increase their fitness (pp.515-16). This led to a review of models and empirical findings supporting the theory that group selection in humans has been led by cultural strategies. Socialization pressure could have included mutual monitoring and moralistic aggression (p.517) that helped shape an optimum degree of ethnocentric ideology that replicated group genes most successfully (p.518).

Limitation of space precludes reviewing all evidence for and against GST as it was applied to ethnicity. Instead we limit discussion to four major objections that are of special relevance to ethnicity and whose failure has left the theory more firmly grounded.

The first criticism is that inclusive fitness processes can only operate between genealogical kin because their genes are identical by common descent (Mealey 1985; Tooby & Cosmides 1989). Hamilton's 1964 paper is cited in support of this objection which, if true, causes the genetic similarity theory of ethnic nepotism to fail. However, Hamilton dropped the identical-by-descent criterion in the early 1970s (e.g. Hamilton, 1996/1971, p.221; discussed by Pepper, 2000, pp.355-6). Genetic similarity is a sufficient basis for inclusive fitness to operate, regardless of how the similarity arises.

A second criticism has been that ethnic kinship is too slight ever to justify diverting effort from genealogical kin. A related criticism is that a gift or other benefit always yields a larger genetic payoff when directed to close kin than to co-ethnics. However, Hamilton himself showed that the aggregate kinship in populations can be sufficient to allow investment in it to be adaptive (Hamilton, 1996/1971, p.221). Harpending derived the same result (Harpending, 1979). And it is feasible for individuals to contribute to the welfare of multitudes via collective goods, such as big game hunting, group defense, and punishment of free-riders (Goetze, 2007).

The third criticism of GST was made by Grafen (1990, p.51) who argued that for altruism to pay off, shared genes must occur frequently in the genome, which he thought was not the case beyond genealogical kin. Citing Grafen, Leek and Smith (1989, p.534) conclude: "[I]t seems important for Rushton to provide a more convincing rationale for the existence of sufficient degrees of overall genetic similarity amongst non-relatives ...". Grafen's criticism may hold for selection within outbred populations. However, ethnic kinship is considerable within ethnically partitioned populations. Ethnic kinship was estimated in 2002 by Harpending to average 15% between unrelated populations (Harpending, 2002). The example given in section 3.2 below finds a kinship of 6%, equivalent to that with a great grandparent, more than sufficient to meet Grafen's criterion.

A fourth criticism proceeded from the reformation in Anthropology that followed World War II and was established in the 1970s consequent to the Vietnam War. According to the dogma people were really gentle, peaceful, and nice except for disruptions caused by colonialism. Even more pervasive was the new "pots not peoples" view of human mobility. In this understanding the earth before agriculture was a carpet of sessile foragers interacting with neighboring groups only over short distances. Given this, people rarely if ever encountered others unlike themselves so there was never any selection in favor of ethnic or racial recognition and discrimination. A familiar saying was that we were foragers "for 99% of our existence". The criticism is dubious for two reasons. First, it discounts the possibility that humans can detect slight phenotypic resemblances. Secondly, the claim of low mobility has difficulties. Agriculture is about 10,000 years old while fully modern humans are only 45,000 or so years old, and we have no reason to think that bursts of innovation and population growth and decline have been limited to agricultural peoples. If, as is likely, human history has always been characterized by eruptions, invasions, and extinctions, then the criticism disappears.

3. Subsequent Genetic Similarity Research on Ethnic Nepotism

3.1 Social behavior

In a follow-up exchange on his 1989 paper, Rushton (1991) hypothesized and discussed methods for detecting ethnic nepotism in science. He recommended studying journal citations to search for assortment of authors. His hypothesis was confirmed using the citation method in a large-scale study of European journals (Greenwald and Schuh, 1994).

In 1997 Rushton discussed the stability of states experiencing changing ethnic proportions. He observed that the Soviet empire had fragmented into constituent nations and argued that ethnic conflict due to genetic dissimilarity was a rising centripetal force within the United States and Canada due to increasing diversity. Welfare was a likely divisive factor (p.375). He examined a proposal for the partial ethnic partition of the US as guided by individual choice as a means for reducing conflict. In a 2005 paper Rushton argued that GST was compatible with the theory of ethno-symbolism, according to which nations form around pre-existing ethnic groups and are perpetuated by the founding ethny's myth-symbol complex.

The theory that ethnic groups are pools of genetic similarity has helped inspire sociological research. GST presaged the study of ethnic altruism as a social force. It conceptualized ethnic groups as extended kinship, with greater genetic similarity within than between. It connected to van den Berghe's concept of ethnic nepotism, which predicts more generous charity within than between ethnic groups (Rushton, 1984, p.78).

Thus from its initial formulation, GST set many parameters of the study of ethnic nepotism. It is consistent with the mainstream definition of ethnicity given earlier, which includes population-level belief in descent from common ancestors. Solidarity is also a characteristic of ethnic groups, again consistent with Weber, and the study of charity and other forms of altruism has successfully tested Rushton's prediction.

Sociological research confirms GST applied to ethnicity (Salter, 2007). Trust and the risky joint enterprises it facilitates are more common within than among ethnic groups (Salter, 2002), giving an ethnic dimension to middleman trading groups, organized crime, political dissidents, and nationalist freedom fighters. Rushton's prediction that ethnic diversity lowers trust and cooperation has been confirmed (Putnam, 2007).

Ethnic similarity also affects charitable donations to strangers in need, as predicted by GST. Field studies of street beggars in Eastern Europe find that passersby give more generously to beggars of the same ethnicity (Butovskaya, Salter, Diakonov & Smirnov, 2000). Global comparisons of government expenditure show that ethno-linguistic diversity explains over 30% of the variance in support for welfare ($r = -0.56$) (Sanderson & Vanhanen, 2004, p.120). Diversity explains 80% of the variance in foreign aid expenditure after controlling for national income and overall budget ($r \approx -0.9$). It also correlates negatively with economic growth for all except the wealthiest 10% of economies (Masters & McMillan, 2004). A likely explanation is that diversity decreases national cohesion and the ability of governments to make rational economic decisions (Alesina, Baqir & Easterly 1999; Alesina & Spolaore 2003). An Australian study has confirmed the similarity-cohesion link (Healy, 2007).

3.2 Genomics and Kin Recognition

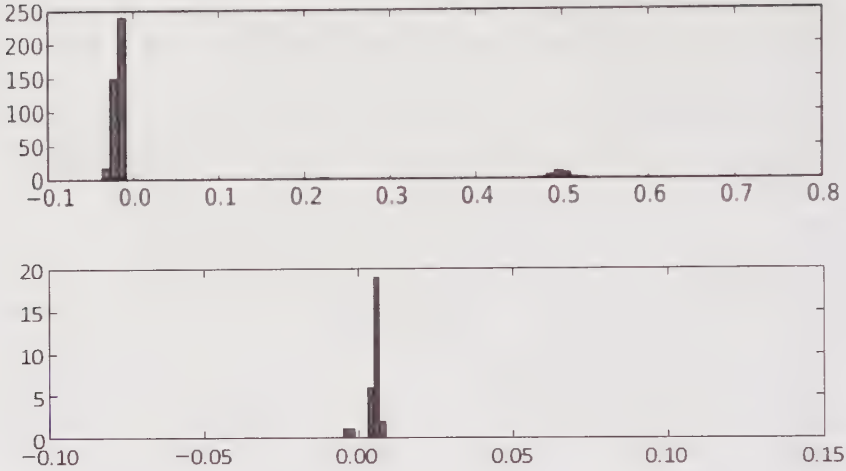
With the arrival of inexpensive technology to genotype individuals at large numbers of single nucleotide polymorphisms we can rather precisely assess genetic similarity (simply "kinship" in the terminology of genetics) between individuals within and between groups. For example worldwide samples of individuals from human populations assessed hundreds of thousands of polymorphic markers are easily available on the internet (http://www.hgdp.org). Here we explore such patterns in several populations. What we present here is essentially an elaboration of previous work that confirmed Rushton's arguments concerning ethnic kinship and the

extra parental kinship resulting from endogamous marriage (Harpending, 1979; 2002; Salter, 2002).

While much of inclusive theory has been developed in terms of the coefficient of relationship, everything is easier when it is written in terms of the coefficient of kinship. For example the coefficient of relationship, the "fraction of shared genes" is unity with oneself. But what if a person is highly inbred? Then we need something to recognize that such a person is "more related" to himself than the offspring of a random mating or an outbred mating. Kinship with oneself in an infinite random mating population is ½ rather than 1, derived like this. Pick an allele from a locus from a person, then pick another from the same locus in the same individual and ask if it is the same. The probability it is the same is just ½, but if the individual is inbred it is greater than ½ and if the individual is outbred it is less. Similar reasoning applies to any pair of individuals. In this formalism, the coefficient of relationship of person *a* to person *b* is the ratio of *a*'s kinship with *b* to *a*'s kinship with himself. This has the strange property that the relationship of *a* to *b* is not necessarily the same as the relationship of *b* to *a* (for details see Harpending, 1979; 2002).

Given the databases published by the Human Genome Diversity project it is relatively straightforward to compute pairwise kinship in their samples. In Figure 1, for example, are results from nearly a million single nucleotide polymorphisms in the 29 individuals of the HGDP French sample. The top panel shows all pairwise kinships while the bottom panel shows, for each individual, how close a kinsman he can find in this sample.

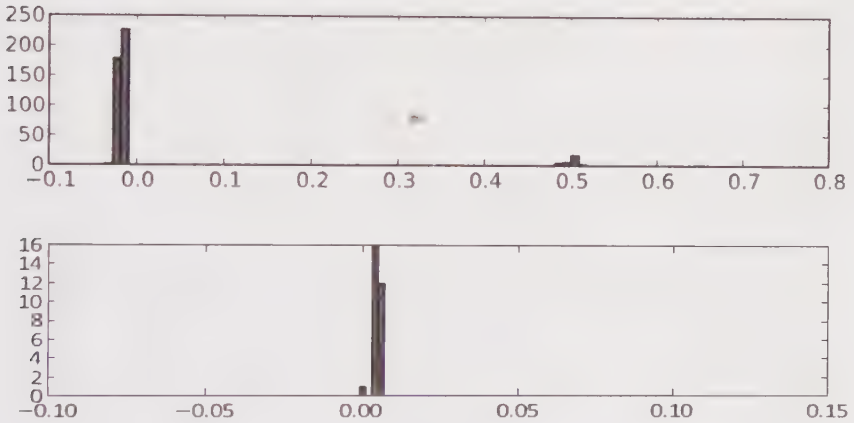
Figure 1. French. The top panel is a histogram of genetic kinship between all possible pairs of individuals in the HGDP French sample, assayed using hundreds of thousands of nuclear genes. The mass at around 0.5 is kinship with self, which varies according to the kinship of the parents with each other. The bottom panel shows, for each individual, his or her closest kin in the sample.



Notice in the top panel of Figure 1 the cluster of mass around ½: these comparisons are kinship with self, showing the variation around the theoretical value of ½ in this large real population. Some are more inbred than others. The bottom panel shows that if these 29 people were, say, adult males in a small community there is little or no opportunity to exploit genetic similarity to form nepotistic cooperative arrangements. The best one can do is around 1%, meaning that helping one's closest (stranger) kin is worth about 2% of helping oneself by the same amount, 4% as much as helping one's own child, etc. Doing so would depend on the ability to discern 1% genetic kinship, which is implausible. Rushton's extensive data showing similarity among friends and spouses is unlikely to be due to such a low degree of genetic resemblance. Altruism via incomplete reciprocity fails for the same reasons. Investing in aggregates is not feasible because strangers are not grouped by family or clan or ethnicity. These genetic data therefore disconfirm Genetic Similarity Theory within outbred populations.

Figure 2 shows the result of the same computation for 29 Japanese.

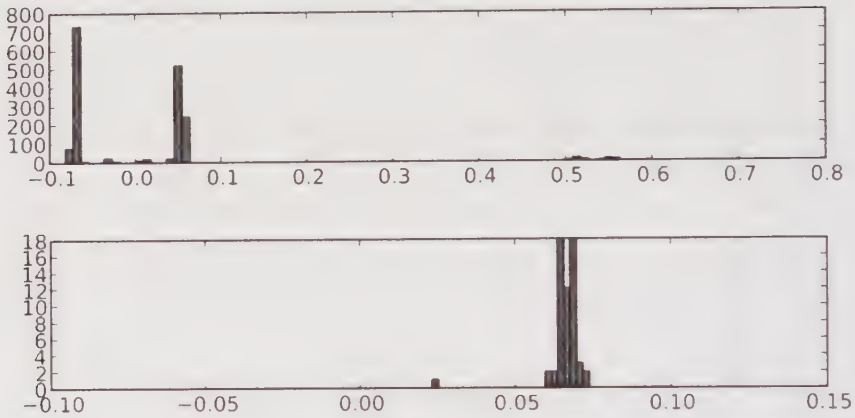
Figure 2. Japanese. The top panel is a histogram of genetic kinship between all possible pairs of individuals in the HGDP Japanese sample, assayed using hundreds of thousands of nuclear genes. The mass at around 0.5 is kinship with self, which varies according to the kinship of the parents with each other. The bottom panel shows, for each individual, his or her closest kin in the sample.



The pattern for the Japanese is hardly different from that of the French. Now let us suppose that these two communities, one of 29 French male adults and one of 29 Japanese adults, were brought together in the one community. In this new community we can repeat the calculations to yield Figure 3.

Figure 3. Mixed. The top panel is a histogram of genetic kinship between all possible pairs of individuals in a synthetic ethnically diverse population created by pooling the French and Japanese HGDP samples. Most people are either positively related to others at about 0.06, equivalent to a great-grandchild, or negatively related to others by the same amount. The bottom panel shows, for each individual, his or her closest kin in the sample. For almost everyone his or her closest genetic kinsman is equivalent to a great-

grandchild.



This new diverse community looks like nature red in tooth and claw in the making. Imagine for example that conditions are Malthusian and that one can share a transient surplus with a neighbor, thereby increasing the latter's individual fitness. If a person can recognize ethnic kin using cultural or heritable markers, he can pick a neighbor with kinship of 0.06 almost every time, corresponding to kinship with a great-grandchild. If at marginal cost he confers some fitness benefit on this neighbor, this is equivalent to increasing his own fitness by 12% (0.06/0.50) of that benefit. On the other hand if he confers the same benefit to a neighbor with kinship -0.06, that decreases his own fitness by the same 12%. Discrimination can therefore cause an action or relationship to yield a 24% difference in fitness. This is an extraordinarily strong selective force, and any quantitative trait that favored ethnic kin discrimination would be rapidly selected with consequences easily visible within a few hundreds to thousands of years. The effect would be stronger when the benefit was conferred on aggregated ethnic kinship via a collective good; so strong as to select for the extreme altruism associated with violent ethnic conflict. Thus GST is confirmed for interactions between populations, such as adjacent or mixed ethnic groups. More

generally, GST is confirmed for situations in which population subdivision creates significant genetic diversity. Another example is endogamous classes, the extreme case being castes. GST would also apply to small isolates, inbred populations where ancestry links are short due to cousin-cousin and uncle-niece marriages. In these populations kinship outside nuclear families is extremely varied, making it possible for altruism between strangers to increase fitness.

3. Amendments and future research

Rushton's theory of ethnic nepotism explains broad trends in ethnic affiliation and has been productive in generating empirical research. One shortcoming in the original formulation is that GST does not account for the way ethnic solidarity varies from culture to culture and within cultures over time. Although solidarity is inherent to ethnicity, it is usually of moderate or low intensity. However, it strengthens in response to attacks perceived to be aimed at group identity, especially invasion of the homeland and physical harm done to co-ethnics. The latter are far more efficient releasers of ethnic sentiment than are perceptions of similarity alone (Salter, 2008a). Even a symbolic threat to group status, such as an ethnic slur, can produce an intense emotional response.

GST needs to be modified if it is to explain both the quiescence and passion of ethnicity. Ethnic nepotism's sensitivity to culture and situation sets it apart from the more constant familial bonds. Sustaining high levels of solidarity requires sustained cultural stimulation. Minorities that have retained ethnic solidarity over long periods have religious rituals that keep alive historical memories of victories and defeats across generations (Spicer, 1971).

An example of the sensitivity of ethnic nepotism to culture and situation is provided by Kurzban, Tooby & Cosmides (2001), who found that participants in an experiment were less prone to categorize others by race when the others' race did not correlate with coalition membership. The same was not true of categorization

by sex, which remained robust even when coalitions contained men and women. This paper might be taken as a challenge to Rushton's theory because the authors concluded that race is a proxy for coalition, a formulation incompatible with ethnic categorization being "automatic and mandatory" and invariably strong. However, this does not apply to GST, which is based largely on intra-ethnic relations and describes weak ties. It should also be noted that Kurzban et al.'s hypothesis might be difficult to generalize because it is based on erroneous evolutionary assumptions. They describe race as a construct that corresponds to inconsequential genetic difference, based on an argument by Lewontin (1972) (p. 15387). Though influential, this argument is fallacious because genetic and phenotypic variation among populations is substantial (Edwards, 2003). Kurzban et al. also assert that in the environment in which humans evolved individuals would not have met members of other races. As noted earlier, this ignores both the human capacity to detect small intra-racial group differences and the high likelihood that regular contact of very different peoples occurred over most of human history.

Amendments have been proposed to both the evolutionary and behavioral components of GST's account of ethnic nepotism, though none that contradict the core insight. As already noted in Section 2, the evolutionary part of the theory, especially involving group selection, becomes more plausible when combined with the concepts of aggregate kinship and collective goods. In competitions between primordial ethnic populations, especially in defending territory, tribal fighters defended aggregate kinships comparable in scale to aggregate family kinship (Salter, 2007/2003, pp.63-7). In the case of hunter-gatherer bands and tribal units it was possible to invest in aggregate kinship by contributing to collective goods, such as defense and big game hunting (Goetze, 2007).

The behavioral component of GST is also being improved by findings from social psychology and sociology, aided by the heuristic advantages of theories that combine evolutionary and behavioural mechanisms. In doing so GST is helping unite the

known causes of ethnic behaviour within a single evolutionary theory.

It will be useful to explore interactions between innate categorization of human kinds, social identification, and attraction to similar others. By age three infants categorize themselves and others into descent groups and apply this to distinguishing races (Hirschfeld, 1996). The attraction of phenotypic similarity is weaker than the ties of ethnic identity. However, similarity cues an implicit state that can position individuals for a transition to explicit ethnic identity (MacDonald, 2008). Evolutionary theories that conceptualize organized religion as group strategies point to links between religiously inspired altruism and group reproductive interests (MacDonald, 1994; Wilson, 2002). Brain scan technology is helping illuminate the mental structures that distinguish implicit and explicit ethnicity (Norton, Mason, Vandello & Biga, 2012). Further work is needed to identify the factors that modulate group-identity and which ethnic markers release greatest affiliative motivation. It should also be useful to distinguish the kinds of affiliation involved, and how they interact with moral sentiments (Salter, 2008b). Rushton's important discoveries concerning genetic ethnic similarity are yet to be fully mined.

References

Alesina, A., Baqir, R. & Easterly W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics 114*(Nov.), 1243-1284.

Alesina, A. & Spolaore, E. (2003). *The size of nations*. Cambridge, MA: MIT Press.

Butovskaya, M., Salter, F. Diakonov, I. & Smirnov, A. (2000). Urban begging and ethnic nepotism in Russia: An ethological pilot study. *Human Nature 11*(2), 157-182.

Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. (1994). *The history and geography of human genes*. Princeton, New Jersey: Princeton University Press.

Edwards, A. W. F. (2003). Human genetic diversity: Lewontin's fallacy. *BioEssays 25*(8), 798-801.

Eibl-Eibesfeldt, I. (1972/1970). *Love and hate: The natural history of behavior patterns*. New York: Holt, Rinehart and Winston.

Eibl-Eibesfeldt, I. (1982). Warfare, man's indoctrinability and group selection. *Ethology (Zeitschrift für Tierpsychologie) 60*, 177-198.

Goetze, D. (2007). Public goods, sharing genes, and the formation of large groups. *Politics and the Life Sciences 26*(2), 7-15.

Grafen, A. (1990). Do animals really recognize kin? *Animal Behaviour 39*, 42-54.

Greenwald, A. G. & Schuh, E. S. (1994). An ethnic bias in scientific citations. *European Journal of Social Psychology 24*(Nov-Dec), 623-639.

Hamilton, W. D. (1964). The genetic evolution of social behavior, parts 1 and 2. *Journal of Theoretical Biology 7*, 1-51.

Hamilton, W. D. (1996/1971). Selection of selfish and altruistic behaviour in some extreme models. In W. D. Hamilton (Ed.), *Narrow roads of gene land. Vol. 1: Evolution of social behaviour* (pp.198-227). Oxford: W. H. Freeman.

Harpending, H. (1979). The population genetics of interactions. *American Naturalist 113*, 622-630.
Harpending, H. (2002). Kinship and population subdivision. *Population and Environment 24*(2), 141-147.
Healy, E. (2007). Ethnic diversity and social cohesion in Melbourne. *People and Place 15*(4), 49-64.
Hirschfeld, L. A. (1996). *Race in the making. Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: MIT Press.
Kurzban, R., Tooby, J. & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences 98*, 15387-15392.
Leek, M. & Smith, P. K. (1989). Phenotypic matching, human altruism, and mate preference [commentary on Rushton 1989]. *Behavioral and Brain Sciences 12*(3), 534-535.
Lewontin, R. C. (1972). The apportionment of human diversity. *Evolutionary Biology 6*, 381-398.
MacDonald, K. (1994). *A people that shall dwell alone: Judaism as a group evolutionary strategy*. Westport, Conn.: Praeger.
MacDonald, K. B. (2008). Effortful control, explicit processing and the regulation of human evolved predispositions. *Psychological Review 115*(4), 1012–1031.
Masters, W. & McMillan, M. (2004). Ethnolinguistic diversity, government, and growth. In F. K. Salter (Ed.), *Welfare, ethnicity, and altruism. New data and evolutionary theory* (pp.123-147). London: Frank Cass.
Mealey, L. (1985). The relationship between social and biological success: a case study of the Mormon religious hierarchy. *Ethology and Sociobiology 6*, 249-257.
Pepper, J. W. (2000). Relatedness in trait group models of social evolution. *Journal of Theoretical Biology 206*, 355-368.
Putnam, R. D. (2007). *E Pluribus Unum*: Diversity and community in the twenty-first century. The 2006 Johan Skytte Prize lecture. *Scandinavian Political Studies 30*, 137-174.

Rushton, J. P. (1984). Group differences, genetic similarity, and the importance of personality traits: Reply to commentators. *Annals of Theoretical Psychology 2*, 73-81.
Rushton, J. P. (1989a). Genetic similarity in male friends. *Ethology and Sociobiology 10*, 361-373.
Rushton, J. P. (1989b). Genetic similarity, human altruism, and group selection. *Behavioral and Brain Sciences 12*, 503-559.
Rushton, J. P. (1991). Ethnic nepotism in science? *Behavioral and Brain Sciences 14*(3), 526-527.
Rushton, J. P. (1997). Will Canada and the United States break up like the Soviet Union? A biopolitical perspective. In A. Somit & S. A. Peterson (Eds.), *Research in biopolitics vol. 5* (pp.373-390). Greenwich, Connecticut: JAI Press.
Rushton, J. P., Russell, R. J. H. & Wells, H. G. (1984). Genetic similarity theory: Beyond kin selection. *Behavior Genetics 14*, 179-193.
Rushton, J. P., Russell, R. J. H. & Wells, H. G. (1985). Personality and genetic similarity theory. *Journal of Social and Biological Structures 8*, 63-86.
Russell, R. J. H. & Rushton, J. P. (1985). Evidence for genetic similarity detection in human marriage. *Ethology and Sociobiology 6*, 183-187.
Salter, F. K. (2002). Estimating ethnic genetic interests: Is it adaptive to resist replacement migration? *Population and Environment 24*(2), 111-140.
Salter, F. K. (2007/2003). *On genetic interests. Family, ethnicity, and humanity in an age of mass migration*. New York: Transaction.
Salter, F. K. (2007). Ethnic nepotism as heuristic: Risky transactions and public altruism. In R. I. M. Dunbar & L. Barrett (Eds.), *Handbook of evolutionary psychology* (pp.54-51). Oxford: Oxford University Press.
Salter, F. K. (2008a). Ethnicity and indoctrination for violence: The efficiency of producing terrorists. In I. A. Karawan, W. McCormack & S. E. Reynolds (Eds.), *Values and violence: Intangible aspects of terrorism* (pp.63-79). New York: Springer.

Salter, F. K. (2008b). Westermarck's altruism: Charity releasers, moral emotions, and the welfare ethic. *Politics and the Life Sciences 27*(2), 28-46.
Sanderson, S. & Vanhanen, T. (2004). Reconciling the differences between Sanderson's and Vanhanen's results. In F. K. Salter (Ed.), *Welfare, ethnicity, and altruism. New data and evolutionary theory* (pp.119-120). London: Frank Cass.
Shaw, R. P. & Wong, Y. (1989). *Genetic seeds of warfare: Evolution, nationalism, and patriotism*. London: Unwin Hyman.
Smith, A. D. (1986). *The ethnic origins of nations*. Oxford: Basil Blackwell.
Spicer, E. H. (1971). Persistent cultural systems. *Science 174*(November), 795-800.
Tooby, J. & Cosmides, L. (1989). Kin selection, genic selection, and information-dependent strategies [comment on Rushton (1989)]. *Behavioral and Brain Sciences 12*(3), 542-544.
van den Berghe, P. L. (1981). *The ethnic phenomenon*. New York: Elsevier.
Weber, M. (1946/1922). The nation. In H. H. Gerth & C. W. Mills (Eds.), *From Max Weber: Essays in sociology* (pp.171-179). New York: Oxford University Press.
Wilson, D. S. (2002). *Darwin's cathedral: The organismic nature of religion*. Chicago: University of Chicago Press.

Web reference
Norton, M. I., Mason, M. F. Vandello, J. A. & Biga, A. (2012). An fMRI investigation of racial paralysis. *Social Cognition and Affective Neuroscience*. Published online 20 Jan. http://scan.oxfordjournals.org/content/early/2012/01/20/scan.nss010.abstract, accessed 20 April 2012. DOI: 10.1093/scan/nss010.

Chapter 10

An Examination of Rushton's Theory of Differences in Penis Length and Circumference and r-*K* Life History Theory in 113 Populations

Richard Lynn
University of Ulster, Coleraine, Northern Ireland, United Kingdom, BT52 1SA

Abstract

Rushton's (1985, 2000) *r-K* life history theory that Mongoloids are the most *K* evolved, Caucasoids somewhat less *K* evolved, and Negroids the least *K* evolved is examined and extended in an analysis of data for erect penis length and circumference in three new data sets. These new data extend Rushton's theory by presenting disaggregated data for penis size for European and North African/South Asian Caucasoids; for East Asian and Southeast Asian Mongoloids; for Inuit and Amerindians and Mestizos, and for thirteen mixed race samples. The results generally confirm and extend Rushton's *r-K* life history theory.

1. Introduction

J. Philippe Rushton (1985, 2000) has advanced a theory of race differences in *r-K* life history. The theory is drawn from biology, in which species are categorized on a continuum running from *r* strategists to *K* strategists; *r* strategists have large numbers of offspring and invest relatively little in them, while *K* strategists have fewer offspring and invest heavily in them by feeding and protecting them during infancy and until they are old enough to look after themselves (Wilson (1975). Fish, amphibians and reptiles are *r* strategists (large numbers of offspring and minimum investment) while mammals are *K* strategists (fewer offspring and greater investment). The *K* strategy is particularly strongly evolved in monkeys, apes and humans. Species that are *K* strategists have a syndrome of characteristics of which the most important are larger brain size, higher intelligence, longer gestation, and a slower rate of maturation in infancy and childhood.

Rushton (2000) has applied *r-K* life history theory to the three major races of *Homo sapiens*: Mongoloids (East Asians), Caucasoids (Europeans, South Asians and North Africans), and Negroids (sub-Saharan Africans). His theory is that East Asians are the most *K* evolved and Negroids the least *K* evolved, while Caucasoids fall intermediate between the two although closer to East Asians. Rushton has supported his theory by documenting that the three races differ in brain size, intelligence, length of gestation, rate of maturation in infancy and childhood, and a number of other variables including penis length and diameter.

Rushton (2000, pp.167-169) reports that penis length and diameter are greatest in Negroids, intermediate in Caucasoids and smallest in Mongoloids. He reports that average erect penises are 4 to 5.5 inches in length and 1.25 inches in diameter among Mongoloids, 5.5 to 6 inches in length and 1.5 inches in diameter among Caucasoids, and 6.25 to 8 inches in length and 2 inches in diameter among Negroids (Rushton & Bogaert, 1987). He reports that in the United States, penises are longer among blacks than among whites (16.36 cm and 15.62 cm respectively) and that

standard 52 mm condoms can fit most Caucasoids but are often too small for Negroids and too large for Mongoloids (Rushton, 2000, p.167).

The greater penis length of Negroids than of Caucasoids appears to have been first noted in the first century AD by the Greek physician Galen (AD 130-201) (Lewis, 1990). This was also observed in the Middle East in mediaeval times and was noted in the *One Thousand and One Nights*, the collection of stories of largely 9th century Persian origin and translated in the 19th century by the British Arabist Richard Burton (1885-8). Several of the stories are about Persian wives who obtain sexual satisfaction with black slaves because these gave greater pleasure on account of the larger size of their penises. Similar accounts appear in ancient Jewish and Greco-Roman sources (Goldenberg, 2003, pp. 190, 369-370 n. 48-51; Thompson, 1989, pp. 6, 107, 210 n. 85).

In the 20th century the large penis size of blacks was documented and illustrated with a number of photographs by Bernatzik (1929) and was confirmed Baker (1974, p.329). These accounts have been of comparisons between Negroids and Caucasoids. There is less evidence for Mongolloids, but Rushton cites A French Army Surgeon (1898/1972) for the observation that Mongoloids have smaller penis length than Causasoids.

Rushton's contributions have been to document the race differences in penis length (Negroids>Causasoids>Mongoloids) more fully, and to formulate an evolutionary theory to account for these differences and integrate them with a number of other variables on an r-K life history continuum. His evolutionary theory is based on the premise that Negroids evolved in equatorial sub-Saharan Africa and that the Caucasoids and Mongoloids evolved from groups that migrated into the colder environments of North Africa, the Middle East, Europe and South Asia and Northeast Asia. This theory is now widely accepted by evolutionary biologists, e.g. Stringer (2011). Rushton proposes that these colder environments were more cognitively demanding and these selected for larger brains and greater intelligence. There is widespread consensus on this thesis, e.g. Kanazawa (2008), Lynn (1991, 2006), and Templer

& Arikawa (2006). Rushton extends the theory of these climatic selection effects further by proposing that colder environments selected for populations that had greater complexity of social organisation achieved by stronger co-operation between males and a reduction of inter-male sexual competitiveness and aggression (Rushton, 2000, p.231). The reason for these adaptations was that in the colder climates men had to co-operate in group hunting to secure food and effective hunting required a greater degree of co-operation and a reduction of inter-male sexual competitiveness and aggression than was required in equatorial latitudes, where plant and insect foods are available throughout the year, there is little need for co-operative group hunting is unnecessary, and a high level of inter-male aggression is adaptive for reproductive success. Testosterone is a determinant of aggression (Book, Starzyk & Quinsey, 2001; Brooks & Reddon, 1996; Dabbs, 2000). Hence, a reduction of aggression and sexual competitiveness between men in the colder climates would have been achieved by a reduction of testosterone, entailing the race differences in testosterone (Negroids>Caucasoids>Mongoloids) that are given in Lynn (1990). The reduction of testosterone had the effect of reducing penis length, for which evidence is given by Widodsky & Greene (1940).

Rushton's theory does not propose that there is a direct causal relationship between intelligence and penis length, such that high intelligence entails a reduction in penis length or greater penis length entails a reduction in intelligence. It is more probable that changes in these two variables evolved independently of each other in response to the climatic environments in which the races evolved. The colder environments of Europe and Northeast Asia selected for larger brains, greater intelligence, a reduction of inter-male aggression, and a reduction of testosterone levels and penis length, entailing positively associations of these variables across the three major races.

An elaboration of Rushton's *r-K* theory is that race differences in penis length may have evolved as an adaptation to population differences in the propensity for infidelity among

human females. It has been proposed by Gallup, Burch, Zappieri, Parvez, Stockwell & Davis (2003) that longer penises may have evolved to displace semen left by other males from the female reproductive tract. Thus, under conditions of double mating (sperm competition) the human penis enables males to substitute their semen from those of their rivals. This theory implies that some selection pressures must have operated to reduce inter-male competition and the propensity for infidelity among females in the races that migrated from equatorial Africa into the more temperate and colder environments of Europe and Northeast Asia. A theory to explain this has been proposed by Miller (1994) as what he calls "paternal investment theory". This states that in the colder environments inter-male competition was reduced as a strategy by which men were able to secure females and replaced by increased provisioning of females with food obtained by hunting: "the colder the climate a population evolved in, the more they should have evolved drives that lead to provisioning" Miller (1994, p. 250). The reduced of inter-male competition was secured by a reduction of testosterone, and this entailed a reduction in penis size. The effect of the colder climates on females would have been that they became dependent on males for provisioning them. This selected for fidelity because males would have been more likely to provision faithful than unfaithful female mates: "a female who has a child by a non-loyal male reduces her chance of catching a provisioning mate" (Miller, 1994. p. 242). Miller argues that his paternal investment theory explains why among Europeans and Northeast Asians men and women are more closely bonded, and marriages and non-marital relationships are more stable than those of Africans.

2. New Data on Race Differences in Penis Length

Rushton has not done any further work on race differences in penis length since 2000, but since this year three new data sets on this issue have appeared. The present paper summarises these and makes four contributions to the evaluation of Rushton's theory. First, we examine how far these new data replicate Rushton's claim

for differences in penis length between Mongoloids, Caucasoids and Negroids. Second, we consider data for penis length for European and North African/South Asian Caucasoids, making a distinction between these two groups that Rushton aggregated into a single group. Third, we consider data for penis length for Northeast Asians (Chinese, Japanese and Koreans) and South East Asians (Indonesians, Thais, Malays, etc), making a distinction between these two groups that have sometimes been aggregated by Rushton as Asians. Fourth, we present data for penis length for Inuit, Amerindians, Mestizos, Hispanics and a number of mixed race populations, none of which were considered by Rushton, and consider how far these are consistent with Rushton's theory.

2a. The first new data set consists of a review of a number of studies of penis length and also of penis circumference in five racial samples carried out by Templer (2002). His results are summarised in Table 1. It will be seen that these confirm Rushton's theory that European Caucasoids (represented by samples from Canada and the United States) have greater stretched penis length than Mongoloids, represented by Japanese (4.38 vs 3.37), and that Negroids represented by Nigerians have greater flaccid penis length and circumference than European Caucasoids (represented by samples from Germany, France, Czechoslovakia and the United States) (3.45 vs 3.37 for length and 3.92 vs 3.83) for circumference. Templer reports that the correlation between penis length and circumference is .43.

In addition, Templer's data extend Rushton's theory by distinguishing between European Caucasoids and South Asian Caucasoids, who are aggregated by Rushton into a single group. His data show that European Caucasoids have greater flaccid penis length and circumference than South Asian Caucasoids represented by a sample from India (3.45 vs 2.87 for length; 3.82 vs 3.53 for circumference). Standard deviations were not given in these studies, so it is not possible to calculate the statistical significance of the differences between the groups.

Table 1. Penis length (inches) in five groups (Source: Templer, 2002)

Race	N	Measure	Length	Circumference
South Asians	111	Flaccid	2.87	3.53
Europeans	483	Flaccid	3.45	3.82
Nigerians	320	Flaccid	3.62	3.92
Europeans	142	Stretched	4.38	-
Japanese	184	Stretched	3.37	-

2b. The second new data set consists of a survey carried out in New York City in 2010 of objectively measured and self-measured erect penis length in a sample of 800 adult men consisting of 200 from each group of blacks, Hispanics, whites and Asians. The majority of the group identified as Asians will have comprised Chinese and Koreans, but may have included some south and southeast Asians). These data are reported in http://www.targetmap.com/viewer.aspx?reportId=3073. Retrieved 20 June, 2011 and are given in Table 2. The results confirm Rushton's theory of differences in penis length by finding that both objectively measured erect penis was greatest in Negroids (6.9 inches), intermediate in Caucasoids (6.1 inches) , and smallest in Mongoloids (5.3 inches). The same differences are present for self-assessed penis length but objectively measured penis length is likely the more reliable result. This study also showed that Hispanics had an objectively measured penis length of 6.5 inches, intermediate between that of Negroids and Caucasoids. Standard deviations were not given, so it is not possible to calculate the statistical significance of the differences between the groups.

Table 2. Penis length (inches) in four ethnic groups in New York

Penis length	Blacks	Hispanics	Whites	Asians
Measured	6.9	6.5	6.1	5.3
Self-reported	7.9	7.7	7.8	6.7

2c. The third new data set consists of a worldwide summary of a number of studies of erect penis length in adult men reported for 113 countries in 2010 in http://www.everyoneweb.com/worldpenissize/ Retrieved 20 June, 2011. The information in this website has been collated from data obtained by research centres and reports worldwide. The references for the studies from which the composite data have been compiled are given in the website. The results of this worldwide summary of studies of penis length are shown in Table 3. This gives data for seven racial groups and for a further group of thirteen mixed race populations. In some of these studies erect penis length was self-reported (denoted by asterisks) and in others it was measured by others. The statistical significance of the differences in penis length between the groups is tested by t tests and Scheffé tests in a series of paired comparisons shown in Table 4. Both tests give the same results.

Table 3. Penis length (cm) in 113 countries

Country	Penis length	Country	Penis length
Negroids		**Caucasoids: European**	
Angola	15.73*	Albania	14.73*
Benin	16.20*	Argentina	14.88
Burkina Faso	15.89*	Armenia	13.22
Cameroon	16.67*	Australia	13.31
Central African Republic	15.33	Austria	14.16
Chad	15.39	Azerbaijan	13.72
Côte d'Ivoire	15.22	Belarus	14.63
Congo (Brazzaville)	17.93*	Belgium	15.85*
Congo (Zaire)	17.33*	Bosnia	15.67*
Gambia	15.88	Bulgaria	15.02
Ghana	17.31*	Canada	13.92
Haiti	16.01	Chile	14.59
Jamaica	16.30*	Croatia	14.77
Nigeria	15.50	Cuba	15.29*
Senegal	15.89	Czech Republic	15.89*
South Africa	15.29*	Denmark	15.89*
Zambia	15.78*	Estonia	13.78*

Country	Penis length
Zimbabwe	15.68
Mean	*16.07*
Caucasoids: North African/South Asian	
Afghanistan	13.69
Algeria	14.19*
Bangladesh	11.20
Egypt	15.69
India	10.24
Iran	14.55*
Iraq	14.55*
Israel	14.38*
Lebanon	16.82*
Libya	13.74
Morocco	15.03*
Pakistan	15.25
Palestine	15.08
Saudi Arabia	13.80*
Sri Lanka	10.89
Tunisia	15.01*
Turkey	14.11
Turkmenistan	13.48
Yemen	12.72
Mean	*13.28*
Mongoloids	
China	10.89
Hong Kong	11.19
Japan	10.92
Korea: North	9.66
Korea: South	9.66
Singapore	11.53
Taiwan	10.78
Mean	*10.66*
South East Asians	
Cambodia	10.04
Indonesia	11.67

Country	Penis length
Finland	13.78*
France	13.53*
Georgia	16.00*
Germany	14.48
Greece	14.73
Hungary	15.61*
Iceland	16.51*
Ireland	12.78
Italy	15.74*
Macedonia	13.98
Netherlands	15.87
New Zealand	13.99
Norway	14.34
Poland	14.29
Portugal	13.19
Romania	12.73
Russia	13.21
Serbia	14.87
Slovakia	15.21*
Slovenia	15.13*
Spain	13.85
Sweden	14.88
Switzerland	14.35
Ukraine	13.97
United Kingdom	13.97
USA	12.90
Uruguay	15.14*
Mean	*14.51*
Amerindians/Mestizos	
Colombia	17.03
Ecuador	17.77
El Salvador	14.88
Honduras	15.00
Mexico	15.10*
Nicaragua	16.26*
Panama	16.27*
Paraguay	15.53
Peru	16.03*

Country	Penis length	Country	Penis length
Malaysia	11.49	*Mean*	*15.99*
Philippines	10.85		
Thailand	10.16	**Mixed Race**	
Vietnam	11.47	Belize	15.75*
Mean	*10.95*	Brazil	16.10
		Cape Verde	14.05
		Costa Rica	15.01
Inuit		Dominican Rep.	15.99*
Greenland	13.87*	Eritrea	14.39*
		Ethiopia	13.53*
		Mongolia	12.77
		Panama	16.27*
		Puerto Rico	16.01*
		Surinam	14.67*
		Sudan	16.47*
		Venezuela	17.03
		Mean	*15.23*

Table 4. Statistical significance of paired comparisons

Group comparisons	T	Scheffé
Negroids – European Caucasoids	.000	.000
Negroids – South Asian/North African Caucasoids	.000	.000
Negroids – Amerindians/Mestizos	.799	1.000
Negroids – Mongoloids	.000	.000
Negroids – South East Asians	.000	.000
European Caucasoids – South Asian/North African Caucasoids	.082	.577
European Caucasoids – South East Asians	.000	.000
European Caucasoids – Amerindians/Mestizos	.000	.022
European Caucasoids – Mongoloids	.000	.000
Mongoloids – South Asian/North African Caucasoids	.000	.000
Mongoloids – Amerindians/Mestizos	.000	.000
Mongoloids – South East Asians	.492	.999
Amerindians/Mestizos – South East Asians	.000	.000
Amerindians/Mestizos – South Asian/North African Caucasoids	.002	.001
South Asian/North African Caucasoids – South East Asians	.000	.000

3. Discussion

The data presented in the present paper confirm and extend Rushton's theory of race differences in penis length. Rushton's theory that penis length is greatest in Negroids, intermediate in Caucasoids and smallest in Mongoloids is confirmed in the three data sets given in Tables 1, 2 and 3. Templer's data given in Table 1 also confirm Rushton's theory by showing that the same differences are present for penis circumference. Rushton's theory is further supported by the data in the New York City study, given in Table 2, and more extensively in the 113 nation study given in Table 3, where penis length is greatest in Negroids (16.07), intermediate in Caucasoids (13.89: the average of European and North African/South Asian Caucasoids), and smallest in Mongoloids (10.66.

Rushton's theory of race differences in penis length is extended by the data presented here for a number of new populations. First, the new data extend Rushton's theory by distinguishing between European Caucasoids and North African/South Asian Caucasoids, who are combined by Rushton into one group. Templer's data given in Table 1 show that European Caucasoids have greater flaccid penis length and circumference than South Asian Caucasoids represented by a sample from India. This result is confirmed in the 113 nation data set given in Table 3, where European Caucasoids have greater penis length (14.51) than North African/South Asian Caucasoids (13.28). Although this difference is not statistically significant, both these and Templer's results are in the opposite direction from what would be predicted from Rushton's theory that when early peoples migrated from sub-Saharan Africa into North Africa and South Asia, and later into Europe and North East Asia, penis length became smaller and intelligence became greater in more cognitively demanding cold winter environments. This theory predicts that intelligence should be greater and penis length should be smaller in European Caucasoids than in South Asian/North African Caucasoids, because Europe has had colder winters than South Asia and North Africa.

This prediction is confirmed for intelligence and brain size given by Rushton (2000, pp. 40, 214, 279-80) and Lynn (2006) but in the present two data sets not for penis lengths. These results are therefore anomalous for Rushton's theory and present a problem for it.

Second, the present data make it possible to distinguish between the penis length of Northeast Asians (Chinese, Japanese and Koreans) and Southeast Asians (Indonesians, Thais, Malays, etc), that Rushton aggregated into a single group as East Asians. It can be predicted from Rushton's theory that Northeast Asians (Mongoloids) should have smaller penis length than Southeast Asians because they evolved in a colder environment. The penis length differences given in Table 3 show that this prediction is confirmed (Northeast Asians: 10.66; Southeast Asians: 10.95), although this difference is not statistically significant.

Third, Rushton does not consider the penis length of Hispanics and Mestizos for which data are given in Tables 2 and 3. Hispanics and Mestizos can be regarded as comparable peoples because while Hispanics are not racially homogeneous, most resemble Mestizos as mixed peoples with European and Native American Indian ancestry. The New York City study shown in Table 2 gives data for penis length for Hispanics compared with European Caucasoids, and the 113 nation study shown in Table 3 gives data for penis length for Mestizos in Latin America compared with European Caucasoids. In both data sets, penis length is greater for Hispanics/Mestizos than for European Caucasoids at 6.5 inches and 6.1 inches, respectively, in the New York City data, and 15.99 cm and 14.51 cm, respectively, in the 113 nation study. Although Rushton does not consider Hispanics or Mestizos in his *r-K* theory, these results are consistent with his theory because Hispanics and Mestizos have lower IQs that Europeans at 89 for Hispanics in the United States calculated in the meta-analysis of Roth, Bevier, Bobko, Switzer & Tyler (2001) and 86 for Mestizos in Latin America calculated by Lynn (2006). These results in the two studies are therefore predictable from Rushton's theory and provide further strengthening for it.

Fourth, Table 3 gives the penis lengths of Latin American Mestizos (15.99) compared with those of the Negroids (16.07). This difference is very small and not statistically significant but is nevertheless in the direction that would be predicted from Rushton's theory because Mestizos have higher average IQs than Negroids given in Lynn (2006) as 67 and 86 respectively, and would therefore be expected to have smaller penis length. It may, however, be regarded as something of an anomaly that this difference is not greater than that given in Table 3.

Fifth, the penis length of the Inuit (13.87) is greater than that of Mongoloids (10.66) and smaller than that of Negroids (16.07). This result is predictable from Rushton's theory because the IQ of 91 of the Inuit is also midway between that of Mongoloids and of Negroids, given as 105 and 67, respectively, adopting the figures given in Lynn (2006).

Sixth, we consider now the thirteen countries that have mixed race populations for which penis length data are given in Table 3. The percentages of the races in populations of these countries are taken from Philips (1996) unless otherwise stated.

In **Belize**, the penis length of 15.75 is about what would be predicted in this mixed race of the population consisting of 55 percent Amerindians and Mestizos (penis length:15.99), 30 percent Mulatto, a mix of European Caucasoid (penis length:14.51) and Negroids (penis length:16.07), the average of which is 15.29.

In **Brazil**, the penis length of 16.10 is virtually the same as that of Amerindians and Mestizos (15.99) and that of Negroids (16.07), who together comprise 47 percent of the population, although it is greater than that of European Caucasoids (14.51), who comprise 53 percent of the population.

In **Cape Verde**, the population consists of 28 percent Negroids and 71 percent Mulattos. Penis length of 14.05 is smaller than that of Negroids (16.07) and Mulattos (15.29). This result is an anomaly for Rushton's theory from which it would be expected that the penis length would be intermediate between that of Negroids (16.07) and Mulattos.

In **Costa Rica**, the population consists of 94 percent European Caucasoids and Mestizos. The penis length (15.01) is intermediate between that of Amerindians and Mestizos (15.99) and European Caucasoids (14.51), so this result is about what would be predicted from Rushton's theory.

In the **Dominican Republic**, the population consists of 73 percent Mulattos, 16 European Caucasoids and 10 percent Negroids. Penis length of 15.99 is greater than that of Mulattos (15.29) and European Caucasoids (14.51) but smaller than that of Negroids (6.07). This result is about what would be predicted from Rushton's theory.

In **Eritrea** and **Ethiopia**, the populations are a mix of Negroids and North Africans (Cavalli-Sforza, Menozzi & Piazza, 1994, p. 191), and the penis lengths of 14.39 and 13.53 are intermediate between those of Negroids (16.07) and North Africans (13.28), as would be predicted from Rushton's theory.

In **Mongolia**, the population consists of a mix of Mongoloid, Kazakh and Kirghiz peoples (Cavalli-Sforza, Menozzi & Piazza, 1994, pp. 223, Fig 4.9.1). Penis length of 12.77 is greater than in any of the seven Mongoloid samples (average 10.66), and smaller than that of the two central Asian peoples represented by Turkmenistan (14.48) as would be predicted from Rushton's theory.

In **Puerto Rico**, 76 percent of the population identify themselves as white, but 53.6 percent have some Amerindian ancestry (Madrigal, 2006, p. 121). There are also significant numbers of predominantly African descent, estimated at about a quarter of the population, so the population is a mix of European Caucasoid, Mestizos and Negroids. Penis length (16.01) is virtually identical to that of Mestizos (15.99), fractionally smaller than that Negroids (16.07) but larger than that of Europeans (14.51), so this result is about what would be predicted from Rushton's theory.

In **Panama**, the population is 64 percent Mestizo, 14 percent Negroid and Mulatto, and 10 percent European Caucasoids. Penis length (16.27) is greater than that of Mestizos (15.99),

Negroids (16.07) and Mulattos (15.29), so this result is anomalous for Rushton's theory.

In **Sudan**, 49 per cent of the population are Muslim northerners identified as "Sudanese Arab" according to Phillips (1996, p.39) and the remainder are Christian/Animist southerners. The term "northerner" includes groups like the Fur and the Nubians who are predominantly black African, and even the term "Arab" includes subgroups that are predominantly black African. It is likely that the data come largely from black African subjects. If this is so, penis length (16.47) is fairly close to that of Negroids (16.07), so this result is consistent with Rushton's theory.

In **Surinam**, the population is 35 percent Negroid and Mulatto, 33 percent South Asian, and 16 percent Indonesian. Penis length (14.67) is intermediate between that of Negroids (16.07), Mulattos (15.29), South Asians (13.28) and Indonesians (11.67), as would be predicted from Rushton's theory.

In **Venezuela**, the population is a mix of European Caucasoids, Amerindians, Mestizos and Negroids. Penis length is high at 17.03 and somewhat greater than would be predicted from Rushton's theory.

4. Conclusions

The data given in Tables 1, 2 and 3 provide three confirmations of Rushton's theory that penis length is greatest in Negroids, smaller in Caucasoids and smallest in Mongoloids. In addition, Rushton's theory has been extended in the data presented here by distinguishing between Northeast Asians and Southeast Asians, and between European Caucasoids and North African/South Asian Caucasoids, and by giving penis lengths for a number of other peoples including Inuit, Mulattos, Amerindians, Mestizos and Hispanics, and for thirteen mixed race populations. For most of these populations, penis lengths are predictable and confirmed. The only seriously problematic anomaly is that penis length was found to be greater in European Caucasoids than in South Asian/North African Caucasoids in the two studies whose results are given in

Tables 2 and 3, while Rushton's theory predicts that it should be smaller. However, this anomaly should not be regarded as fatal to Rushton's theory. A possible explanation for it is that penis length is positively related to height at a correlation of .45 in as study by Fisher (1964), confirmed by Siminoski & Bain, (1988), and European Caucasoids are typically taller than South Asian/North African Caucasoids. There are some minor anomalies in the results for the mixed race populations of Cape Verde and Venezuela, but these can be regarded as sampling errors.

Considering the results as a whole in terms of Popper's (1959) theory of the logic of scientific explanation, Rushton's theory has survived three attempts at falsification, all of which have failed, and has generated predictions for penis lengths in other races and populations most of which have been verified. In both of these respects, Rushton's theory can be regarded as a progressive research program that has been strengthened by the data presented in this paper.

References

A French Army Surgeon (1898/1972). *Untrodden fields of anthropology*. Paris: Carrington (Reprinted Huntington, New York: Krieger).

Baker, J.R. (1974). *Race*. Oxford, UK: Oxford University Press.

Book, A. S., Starzyk, K. B. & Quinsey, V. L. (2001). The relationship between testosterone and aggression: A meta-analysis. *Aggression and Violent Behavior*, *6*, 579-591.

Bernatzik, H.A. (1929). *Zwischen weissem Nil und belgisch-Kongo*. Wien: Seidel.

Brooks, J. H. & Reddon, J. R. (1996). Serum testosterone in violent and non-violent young offenders. *Journal of Clinical Psychology*, *52*, 475-483.

Burton, R.F. (1885-8). *The Book of The Thousand Nights and a Night*. London: The Burton Club.

Cavalli-Sforza, L.L., Menozzi, P. & Piazza, A. (1994). *The history and geography of human genes*. Princeton, NJ: Princeton University Press.

Dabbs, J. M. (2000). *Heroes, rogues and lovers: Testosterone and behavior*. New York: McGraw-Hill.

Fisher, L.G. (1964). Penis length and body height. *Medical Aspects of Human Sexuality*, 67, 103.

Gallup, G. G., Jr., Burch, R. L., Zappieri, M. L., Parvez, R. A., Stockwell, M. L. & Davis, J. A. (2003). The human penis as a semen displacement device. E*volution and Human Behavior*, 24, 277-289.

Goldenberg, D.M. (2003). *The Curse of Ham. Race and Slavery in Early Judaism, Christianity, and Islam.* Princeton, NJ: Princeton UniversityPress.

Kanazawa, S. (2008). Temperature and evolutionary novelty as forces behind the evolution of general intelligence. *Intelligence, 36,* 99-108.

Lewis, B. (1990). *Race and Slavery in the Middle East*. New York: Oxford University Press.
Lynn, R. (1990). Testosterone and gonadotrophin levels and *r/K* reproductive strategies. *Psychological Reports*, 67, 1203-1206.
Lynn, R. (1991). The evolution of racial differences in intelligence. *The Mankind Quarterly*, *32*, 99-121.
Lynn, R (2006). *Race differences in intelligence: An evolutionary analysis*. Augusta, GA: Washington Summit Publishers.
Madrigal, L. (2006). *Human biology of Afro-Caribbean populations*. Cambridge: Cambridge University Press.
Miller, E.M. (1994). Paternal provisioning versus mate seeking in human populations. *Personality and Individual Differences, 17,* 227-255.
Nobile, P. (1982). Penis size: The difference between blacks and whites. *Forum: International Journal of Human Relations*, 11, 21-28.
Philips (1996). *World Atlas*. London: Chancellor.
Popper, K.R. (1959). *The Logic of Scientific Discovery*. London: Methuen.
Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: a meta-analysis. *Personnel Psychology*, 54, 297–330.
Rushton, J. P. (1985). Differential *K* theory: The sociobiology of individual and group differences. *Personality and Individual Differences, 6*, 441-452.
Rushton, J.P. (2000). *Race, evolution and behavior*. Third edition. Port Huron, MI: Charles Darwin Research Institute.
Rushton, J.P & Bogaert, A.F. (1987). Race differences in sexual behaviour: Testing an evolutionary hypothesis. *Journal of Research in Personality*, 21, 529-551.
Siminoski, K. & Bain, J. (1988). The relationships among height, penile length, and foot size. *Annals of Sex Research, 6,* 231.
Stringer, G. (2011). *Origin of our Species*. London: Allen Lane.
Templer, D. I. (2002). *Is size important?* Pittsburgh, PA: CeShore.

Templer, D. I. & Arikawa, H. (2006). Temperature, skin color, per capita income, and IQ: An international perspective. *Intelligence, 34*,121–139.

Thompson, L.A. (1989). *Romans and B*lacks. Norman and London: University of Oklahoma Press.

Widodsky, H.S. & Greene, R.R. (1940). The effects of testosterone, estrone, and estrodiol applied locally to the penis of the rat. *Endinocrinology,26*,1078-1080.

Wilson, E. O. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Harvard University Press.

Chapter 11

Migratory Selection for Inversely Related Covariant T- and IQ-Nexus Traits: Testing the IQ/T-Geo-Climatic-Origin Theory by the General Trait Covariance Model

Helmuth Nyborg
University of Aarhus, Denmark (1968-2007)

Abstract

The IQ/T-*Geo-Climatic Origin (GCO)* theory admits that primordial northbound migration out of Africa increasingly demanded higher IQ and decreased aggression as temperatures sank. To organisms with a fixed intra-systemic energy budget, this meant that small-brained – High T (Lo-IQ/Hi-T)" masculine aggressive A5 males became too "expensive" in cold eco-niches, and were replaced by Low Testosterone androgyne A1 (Hi-IQ/Lo-T) light-weight males with sufficient energy for developing the large energy-greedy brain needed for survival during cold winters. The physiological re-balancing of T-Nexus and IQ-Nexus traits probably occurred during migration over the past 40,000 years.

The moderate heritability of the traits leads us to expect that rudiments of this evolutionary progression can still be identified in contemporary geographic race and androtype distributions of inversely related T- and IQ-Nexus traits. Rushton's three-racial ranking of Life History traits (Table 1 in Nyborg, b, this issue) confirms this, as do analyses using the General Trait Covariance model on data for five races and pentile IQ bands.

The evolutionary progression seems time-limited, however, as recent decennia witness a decaying Western civilization.

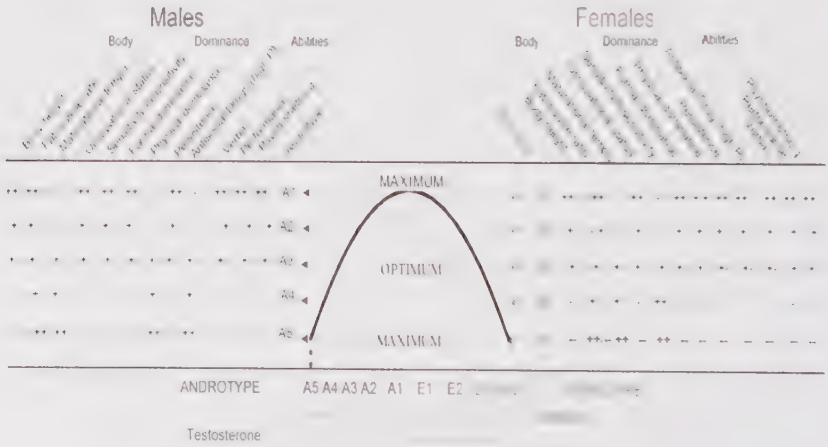
1. Introduction

This paper presents the theory that evolutionary progression took place during primordial northbound migration out of Africa, culminating in Western democracy and civilization. To examine it, we tested a basic premise of the General Trait Covariance (GTC) model (Nyborg, 1987; 1994), and then inspected predictions of the IQ/T-*Geo-Climatic Origin (GCO)* theory.

2. The GTC model

Figure 1 illustrates one of several GTC models

Figure 1. The General Trait Covariance (GTC) model (Nyborg, 1987; 1994) for inverse relationships among General Intelligence-related traits (g-nexus traits) and Testosterone related traits (T-nexus traits) (Female side of model not considered here).



The model generates testable predictions about harmonized body, brain, intellectual, and personality development based on

parental DNA, Testosterone/Estradiol (T/E2) balance, and experience. Optimum brain, intellectual and personality development depends on low and balanced hormone concentrations, at the cost of sexual differentiation, in accordance with the economy principle (Nyborg, 1994). Conversely, maximum sexual differentiation accompanies high and contrasting T and E2 concentrations, respectively, at the cost of optimum intellectual and personality development – the inverse of Hi-IQ/Lo-T relationships (Nyborg, ibid. The present study focuses on the male side of the model to study evolutionary progressive diversion.

2.1. Testing inverse IQ/T-relationships

The model depicts an inverse relationship between IQ (large brains) and T(low aggression), so we first have to provide empirical support for this.

2.1.1. Data and methods.

The large-scale Vietnam Era Study (VES, 1989) provides relevant data for representative groups of non-Hispanic white (N=3,654), Hispanic (N=200), black (N=525), Asian (34), and Native (49) middle-aged American males.

The cognitive test battery consists of 19 experimentally independent variables, highly diverse in types of abilities, information content, and cognitive skills. Five of the tests were administered at the time the subjects were inducted into the armed forces; all the others were administered approximately 17 years after induction, on average. The full battery of 19 cognitive tests is described in details by Nyborg and Jensen (2000a; b). Raw scores were subjected to Principal Component (PC) analyses to extract *g*-factor scores and then converting them to IQ metrics.

Plasma T values were determined at on average 38.1 years of age. Blood specimens were taken in the morning before breakfast, following an overnight fast commencing at 7 p.m. Plasma T concentration was determined (nanograms/deciliter, or n/100ml) using a standard double antibody radioimmunoassay system (Leeco Diagnostics, Inc), and monitored with bench and blind repeat quality control procedures. Vietnam and Non-Vietnam veterans did not differ in T concentration, so data were pooled.

Formal education, income in 1985/86 US dollars, 14 MMPI special scale scores and 4 MMPI-derived Eysenckian personality dimensions were also noted.

2.1.2. Problem.

Ellis and Nyborg (1992) previously found in the VES material that "Old" black T at age 38.1 was only 3.3% higher than non-Hispanic white T. Even if significant, the difference is small, and a previous study found that young black Americans have 19% higher T than young white males (Ross et al., 1986).

The problem is that T drops differently with age for the races. Table 1 provides racial regressions coefficients for the differential T decline with age between 31-49 years.

*Table 1. Regression coefficients used to define Young T values with age 25 (T25) according to the formula[1]: T(25) = (age at measurement – 25) * race tilt coefficient + T at time of measurement.*

Race	Race regression tilt coefficients	Correlations and Level	
Black	y = 1686.89 – 26.31*x	r = -0.28 p<0.001	r^2 = 0.08
Hispanic	y = 1348.73 – 17.80*x	r = -0.17 p<0.01	r^2 = 0.03
White	y = 1243.32 – 15.16*x	r = -0.16 p<0.001	r^2 = 0.03
Asian	y = 1030.98 – 8.15*x	r =-0.08 n.s.	r^2 = 0.01
Native	y = 817.36 – 3.78*x	r = -0.05 n.s.	r^2 = 0.00
All	y = 1296.07 – 16.47*x	r = -0.18 p<0.001	r^2 = 0.03

Note[1]: Thomas Lill Madsen, Tryg Insurance Company, Copenhagen, kindly suggested and performed this analysis, which is gratefully acknowledged.

Black, Hispanic, and white T declines significantly, whereas Asian and Native T does not.

Regression coefficients were used to estimate "Young" T(25)-values from measured Old T(38.1) at mean age 38.1, using the formula in Table 2.

T(38.1) seriously underestimates race differences in T(25) – most for "southernmost" Afro-American blacks, less for "intermediate" Hispanics and "northern" whites, and least for "northernmost" Asians and their Native brothers, who later

migrated to the Americas. Linear retro-estimation undoubtedly inflates T(25) values somewhat, but the actual function is not known.. The North-South reference implied is based on the fact that skin reflection correlates -0.92 with IQ (Templer, in this issue).

Table 2. Retro-estimated Young Testosterone, T(25) at age 25, and measured Old T and Old IQ at age 38.1, by race and sorted by ascending retro-estimated T(25) values.

		T(25)		T(38.1)		IQ(38.1)	
Race	**N**	**Means**	**SD**	**Means**	**SD**	**Means**	**SD**
Native	49	722.82	255.99	678.06	256.25	97.21	14.54
Asian	34	827.20	254.38	727.71	255.09	100.58	16.44
White	3654	864.35	227.35	675.96	230.48	102.57	14.02
Hispanic	200	903.75	256.05	680.72	260.01	91.89	12.67
Black	525	1029.17	238.61	701.43	248.45	84.95	12.32
All groups	4321	883.67	237.16	679.59	234.60	100.00	15.00
F =		65.52; $p<0.001$		1.72; n.s.		196.05; $p<0.001$	
Pearson r with IQ(38.1)		-0.13; t=-8.56; $p<0.001$		-0.07; t=-4.61; $p<0.001$			
Spearman r_s with IQ(38.1)		-0.12; t=-8.00; $p<0.001$		-0.07; t=-4.32; $p<0.001$			

Clearly, studies relying on T(38.1), measured few years before racial slope lines approach each other, will underestimate differences in T(25), as illustrated in Figure 2.

Interestingly, individual differences in T are about 50% heritable (Hoekstra, Bartels & Boomsma, 2006), but most race differences in plasma T become negligent at age 40, even if they affect health (Ellis & Nyborg, 1992) and still exist at very old age (Orwoll et al., 2006).

The small Natives and Asians samples (N=49 and 34, respectively) compromise their IQ means, and Lynn and Vanhanen (2006) suggest that Asian mean IQ is 105. This aberration reduces the true strength of the inverse relationship between T and IQ, but

both Pearson and Spearman correlations nevertheless became significant (Table 2).

Figure 2. Mean differences in retro-projected Young and Middle-age plasma testosterone T, according to 2 age- and five race-categories. Overall Wilks lambda=0.42, F(8, 8,912)=597.13, p<0,001. Univariate Race results: T(25) F=65.52, p<0.001; T(38.1) F=1.72, n.s. Vertical bars denote 0.95 confidence intervals.

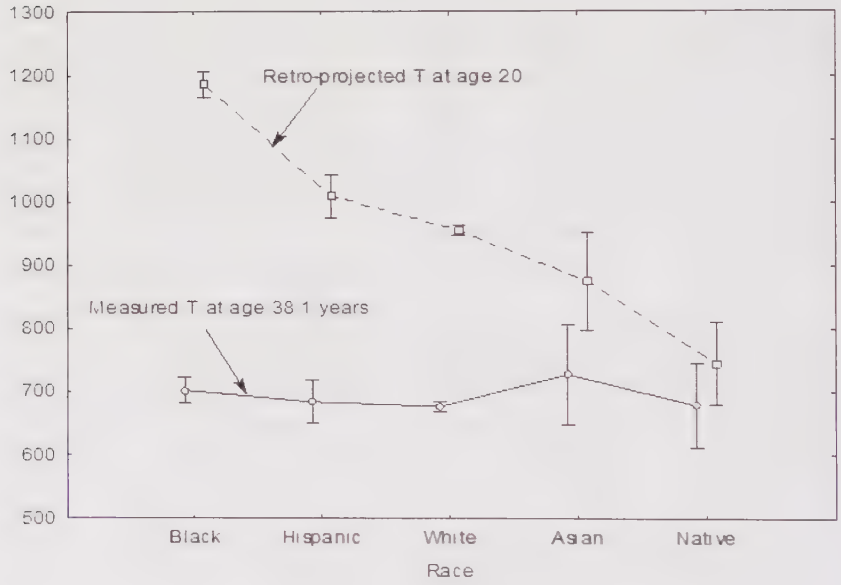


Table 3 provides average T(25), T(38.1), and mean IQ, as categorized by pentile IQ bands

Table 3. Testosterone by pentile IQ bands.

Pentile IQ band	N	T(25)		T(38.1)		IQ(38.1)	
		Means	SD	Means	SD	Means	SD
Very High	373	832.25	207.63	635.26	211.31	124.42	3.61
High	856	851.27	217.36	658.19	217.88	114.61	2.90

Pentile IQ band	N	T(25) Means	T(25) SD	T(38.1) Means	T(38.1) SD	IQ(38.1) Means	IQ(38.1) SD
Average	1944	886.58	234.53	691.85	234.51	100.40	5.65
Low	686	911.79	257.42	695.48	252.56	85.23	2.85
Very Low	462	935.93	259.82	688.00	251.68	73.47	5.09
All Groups	4321	884.17	237.43	680.46	235.00	100.00	15.00
F =		16.67; $p<0.001$		7.38; $p<0.001$			

T-differences are significant across pentile IQ bands.

Figure 3. Mean differences in retro-projected Young and Middle-age plasma testosterone T, according to 2 age- and five pentile IQ bands. Overall Wilks lambda=0.92, F(8, 8630)=45.54, p<0.001. Univariate Race results: T(25) F=16.67, p<0.001; T(38.1) F=7.38, p<0.001. Vertical bars denote 0.95 confidence intervals.

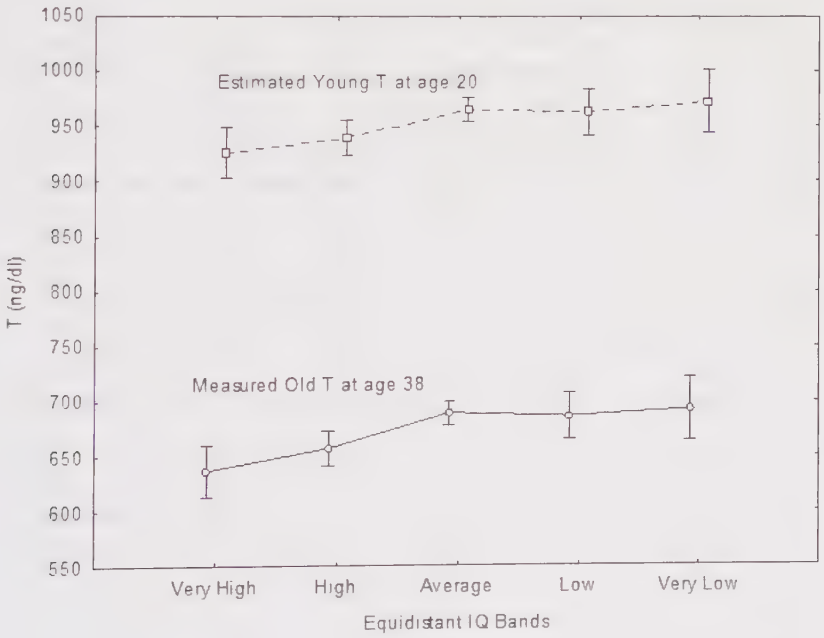


Figure 3 shows that T(25) relates inversely and linearly to IQ, whereas T(38.1) only do so for above average IQs.

. Having addressed the ***how*** question we may now proceed to examine the ***why*** and ***when*** questions of progressive differentiation in evolution.

3. The Inverse IQ/T-Geo-Climatic Origin (GCO) theory

GCO theory capitalizes on an idea recently catching momentum (e.g. Kanazava, 2008; Lynn, 1991; 2006; Nyborg, 1987; 1994; 2012a; Rushton, 1995; 2000; Templer & Arikawa, 2006) that North-bound migration out of Africa 195.000-40.000 years ago through still col der geo-climatic zones selected for large-brained individuals high in IQ and low in aggression.

*3.1. **Why?***

I previously defined the term "Universal Darwinian selection" as selection at the mass-molecular level for the most economic systems, including the "life" and "death" of organic and inorganic molecular constellations, such as men or stars (Nyborg, 1994, pp. 25-27, chapter 10 and 13) – basically in terms of trade-offs among energy "expenses" within fixed budgets. GCO theory draw upon Universal Darwinian selection to explain ***why*** northbound migration gradually had to alter the intra-systemic energy balance to survive: As temperatures sank, individuals were increasingly subjected to selective pressures working with two suites of inversely related covariant physiological body-, brain-, and behavioral traits. Survival depended on a redistribution of energy trade-offs via proximal physiological mechanisms as an obligatory intra-systemic response to increasingly colder environments.

Northbound adaptation thus required a reversal of the original African intra-systemic energy trade-off, so that still less energy became devoted to now maladaptive energy-greedy fetal and pubertal masculinization of body, brain, and behavior, leaving more energy free for developing and maintaining a larger energy-greedy physico-chemical brain. In molecular terms, this translates into a larger physic-chemical brain and still lower levels of masculinizing plasma testosterone (T) as migrants went north.

3.2. When.

The ***when*** question may be considered in terms of the start, transitional, and end conditions for northbound migration.

3.2.1. The origin.

The origin of Man refers to a long evolutionary process selected for still larger and more efficient brains, but with remaining "room at the top" (Rushton & Rushton, 2004). This linear process probably ran slowly in tropical areas, where a favorable climate made most essential animal and plant food resources readily available year round – aside from periodic severe famine. However, the absence of a collective modern protective law system put evolutionary premium on a strong aggressive physique, capable of conquering and defending territory, food, and females. Moreover, an early maturation and high fertility was required to counter the massive number of lives wasted through devastating tribe fights and an astronomic child mortality rate.

The point is that the African condition provided directive selection for increased brain capacity and IQ-Nexus traits (Jensen, 1998), but relatively speaking stronger selection for T-Nexus traits, such as aggression, reproduction and muscular force. According to the GTC model, the evolutionary outcome is a Lo-IQ/Hi-T A5 male, because simultaneous development of a large brain and a physically strong and aggressive body would have transcended the fixed intra-systemic energy budget (The Economy Principle, Nyborg, 1994, chapter 13). A vestige of this early energy trade-off solution may be relatively strong average African or Afro-American A4-5 phenotypes, still providing advantages in fast energy demanding sports like sprint and boxing, which typically is traded-off by below average IQ and educational achievement. The GTC model demands that covariant IQ-Nexus traits will be inversely related to covariant T-Nexus traits.

3.2.2. The demographic transition.

Northbound migratory exodus from Africa gradually altered the IQ-aggression balance. Colder climates generally favor large brains for survival, as migrants has to invent entirely new solutions for coordinated hunting, clothing, housing, and collection

and preservation of food during long unforgiving winters. The widespread, sparse and yearly alternating nature of resources made only those survive who readily found and willingly shared food and shelter with kinship. IQ, altruism, low T and reduced aggression became increasingly privileged.

Lynn and Vanhanen (2006) and Lynn (2006) have already demonstrated the significant South-North gradient for IQ, thus providing strong support for GCO theory. Regrettably, supporting evidence for the T-Nexus part is weaker (see later).

3.2.3. The end point.

The aggression needed for success in competitive Sub-Saharan Africa became a hindrance in colder areas, where adaptation rather called for individual emotional and physical restraint. Sharing and altruism became favored traits, and the later establishments of complex agricultural and industrial cooperatives required and reinforced the trend toward reduced aggression and selfishness. Aggressive fast reproducing migrant became increasingly punished by fellows and by their fixed intra-systemic energy budget, whereas migrants displaying adaptive restraint, altruistic behavior, and capacity for creative-abstract-idealistic solutions to problems in the cold became privileged.

These geo-climatic changes in evolutionary pressures address the ***when*** question.

4. Testing predictions.

The most logical way to test the physiologically anchored predictions of the General Trait Covariance model, and the predictions of the IQ/T-Geo-Climatic Origin theory, in terms of evolution, would be to compare African populations with populations at the last stop on the northbound migratory path, namely, in arctic areas and in North-East Asia. Unfortunately, there are no comparative studies of these populations in which hormones, IQ, education, income, personality including altruism, and psychopathology, were measured concurrently. Absent this, the present analysis takes data from the Vietnam Era Study (VES,

1989) and tests model and theoretical predictions for 5 races and pentile IQ bands.

Table 4. Testing predictions of the GTC model. Aggregated data from the VES (1989) study. Pentile IQ bands are categorical dividers with the inverse IQ/T- relationships implied.

	Progressive inverse IQ - T trade-offs during north-bound migration												
	Hi-IQ ← → Hi-T												
	Lo-T ← → Lo-IQ												
Suggested Racial Androtype position	Asian		Natives		White		Hispanic		Afro-American		African Black		
General Androtype	A1		A2		A3		A4		A5		All Groups[4]		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	N
Pentile IQ bands	124.42[1]	3.61	114.61	2.90	100.40	5.65	88.23	2.86	78.47	5.09	100.00	15.00	4321
Young Testosterone (ng/dl)[1]	921.77	209.47	941.78	218.17	960.83	238.40	1016.27	268.72	1051.07	270.66	981.23	245.24	4321
Formal Education (Yrs.)	15.75	1.94	14.56	2.08	13.13	1.93	13.13	1.90	11.42	1.92	13.30	2.29	4319
Income (1985 US Dollars)	36849	14054	33854	14918	29197	13468	23395	12875	16141	11578	26953	14567	4258
Eysenckian Personality[2]													
Psychoticism (P-scale)	6.38	2.82	6.58	2.80	6.95	3.11	7.80	3.82	8.74	4.54	7.15	3.40	4321
Extraversion (E-scale)	14.10	4.46	14.60	4.42	14.65	4.65	15.21	4.58	15.39	4.23	14.75	4.56	4321
Neuroticism (N-scale)	11.64	4.83	12.39	5.15	13.17	4.84	13.60	5.14	14.09	4.86	13.05	5.00	4321
Social Desirability (L-scale)	7.48	2.32	7.57	2.31	7.04	2.23	9.20	2.57	8.28	2.55	7.93	2.40	4321
MMPI Special Scales[2]													
Ego Strength (ES-scale)	59.49	8.22	57.12	8.65	53.64	9.60	48.49	11.25	44.18	11.87	53.01	10.77	4321
Masculinity-Femininity (MF-scale)	63.41	9.22	61.50	9.99	58.64	9.14	56.20	9.09	54.86	8.71	58.78	9.29	4321
Maintain Control (K-scale)	58.07	8.34	56.39	9.08	53.86	9.18	51.87	9.60	49.60	9.92	53.89	9.54	4321
Confusion (F-scale)	53.64	7.91	54.37	8.00	56.13	9.28	58.88	11.61	62.28	13.61	56.66	10.21	4321
Hypochondria (HS-scale)	52.98	9.25	53.69	10.45	55.77	11.59	59.01	13.75	61.79	16.20	56.33	12.41	4321
Evade Answering (L-scale)	48.37	6.78	49.07	6.26	50.40	7.08	52.19	7.96	53.62	8.57	50.61	7.38	4321
Schizophrenia (SC-scale)	55.92	11.26	56.29	12.52	57.60	13.61	61.65	16.87	65.65	19.26	58.70	14.92	4321
Depression (D-scale)	57.43	10.55	58.49	13.17	60.84	13.67	63.80	15.34	67.40	15.30	61.25	14.10	4321
Obsessive-Compulsive (PT-scale)	55.79	9.46	56.80	11.35	58.70	12.21	61.55	14.11	63.64	14.25	59.08	13.08	4321
Social Introversion (SI-scale	51.49	9.91	52.39	11.29	54.67	10.98	56.30	10.75	57.89	9.74	54.55	10.93	4321
Hypomania (MA-scale)	56.22	9.95	56.16	10.03	57.05	10.97	60.00	11.87	60.78	12.07	57.67	11.10	4321
Psychopathic Personality (PD-scale)	59.32	10.72	59.82	10.89	60.62	11.54	61.67	12.18	61.75	12.55	60.63	11.58	4321
Paranoia (PA-scale)	56.35	8.56	56.30	9.00	56.85	10.17	57.09	11.73	57.90	13.08	56.85	10.57	4321
Hysteria (HY scale)	57.27	7.73	57.75	8.35	57.45	9.08	57.89	10.61	57.95	12.27	57.62	9.49	4321

Notes. [1] Estimated values; [2]Derived ad modum Gentry, Wakefield & Friedman (1985); [3]Rank ordered by F-value; [4]All group diff.:$p<0.001$ except for last two rows; Pearson correlations, *rs*, are all above 0.80 and Spearman $r_s s$ above 0.90 in the aggregated data set. IQ correlates positively with Years of Formal Education, Income and Ego-, M-F-, and Maintain Control scores, and negatively with the

remaining MMPI scores; Testosterone correlations show the exact opposite pattern.

4.1 Testing predictions of the GTC model and the GCO theory

The GTC model and the GCO theory predict trait patterns consistent with Rushton's Life History theory (see Table 1 in Nyborg, b, in this issue). Inversely related IQ and T-Nexus traits are expected to be distributed along a geo-climatic latitude gradient, with a preponderance of A5s near warm Equator and more A1s situated in cold northern countries.

We thus expect high IQ, low T, long formal education, and formal over physical dominance, high income, emotional sensitivity and persistence, increased altruism and reduced psychopathology in northern Hi-IQ/Lo-T Androtype A1s, and the reverse pattern in southern Hi-T/Lo-IQ A5s.

Table 4 confirms this.

A1s combine high IQ with low T, low scores on Eysenckian personality dimensions, high "Feminine Values", "Ego Strength" and "Maintain Control" scores in the MMPI, and low "Confusion", "Hypochondria", "Evade Answering", "Schizophrenia", "Depression", "Obsessive-Compulsive", "Social Introversion", "Hypomania" and "Psychopatic Personality" scale scores. The "Paranoia" and "Hysteria" scores did not differ among Androtypes.

The following account, modified from Dahlstrom et al. (1972), illustrates the contrast between A1s and A5s :

> Lo- T/Hi-IQ Androtype A1 males " *... tend to display feminine values, attitudes and interests, and in styles of expression and speech, as well as in sexual relationships*". A1s are " *... sensitive and prone to worry, idealistic and peaceable, sociable and curious, have general aesthetic interests, and are altruistic.*"
>
> In psychopathological contrast, the Hi-T/Lo-IQ Androtype A5 male " *... tend to display pessimism of outlook on life and the future, feelings of hopelessness and worthlessness, slowing of thought and action, and frequently are preoccupied with death and suicide. He*

... reveal[s] neurotic or moderate psychotic reactions or behavioral disturbances that affect test cooperation. He experiences social uneasiness in social situations or in dealing with others, attempt to deny impulses, temptations, and mental aberrations, and show strong self-depreciatory trends."

All this dovetails nicely with Rushton's (Table 1 in Nyborg, b, in this issue), data on Life history traits, where *K* relate positively to IQ and negatively to T. Table 4 demonstrates that A1s (mostly North-East Asian and non-Hispanic white *K*-strategists, with lesser percentages of other races) also get better educated and earn more than A5s (mostly Africans, but again with lesser percentages of other races).

A1s thus define the culmination of a diversifying evolutionary progression – presaged by GCO theory and illustrated in the GTC model. Importantly, however, androtyping is not race-limited but refers to hormonal continua within and across races.

4. Discussion

The analysis illustrated several points:

IQ and T are inversely correlated across race and pentile IQ bands.

Rushton's racial ordering of 60+ traits (see Nyborg, b, Table 1, in this issue), includes brain size and dovetails nicely with the A1-A5 continuum as racially categorized geographically anchored aggregated data.

In terms of IQ- and T-Nexae, A1s display the predicted suite of traits characterizing high IQ-Nexus individuals, including altruism, reduced masculinity and low psychopathology. A5s display the opposite suite of traits, including increased psychopathology –as predicted by Rushton (2000).

The analysis suggests that altruism is essentially a high IQ – low T phenomenon, and that the contrasting southernmost A5 type reflects remnants of Sub-Saharan existence long before harsh

northern geo-climatic selection took place. A1 is the culmination of harsh selection in the cold.

The retro-calculated significant race differences in young T(25) almost disappear at age 40. Dark-skinned young people have higher T at age 20 than light-skinned individuals, but their T level drops at a faster rate. Given that T is related to crime, the decline might partly, certainly not fully, explain the observed reduction in crime rates with age, and perhaps even better so for blacks.

The obligatory drop in T makes it understandable why older males may act more altruistic and caring than when they were young, reckless, or just careless. The differential T decline rates lead us expect that old blacks and whites would be differentially affected.

The key to understand how the inverse trait covariance Life History trait pattern became established during migration lies arguably in the intra-systemic energy distribution. The modern A1 and ancient A5 types contrast, because prehistoric aggression-related T became selected against, as North-bound migrants proceeded through still colder geo-climatic areas. This left intra-systemic energy free for developing still larger energy-greedy brains, capable in the end of displaying high IQ, altruistic behavior and idealistic attitudes and other traits, not diverted by high T.

Theoretically speaking, androtyping complements Rushton's gene-centered view of the evolution of altruism and other traits, according to which genes for altruism are favored if altruism is directed towards kinship which share the putative genes (i.e. the gene-centered view). GCO theory adds that cold climates favored Lo-T/Hi-IQ A1 altruists via interacting gene-hormone molecular mechanisms during Universal Darwinian assorting. This is conceptualized in the GTC model as a trade-off between an energy-greedy aggressive and reproductively hyper-active A5 Androtype and an equally energy-greedy, but physically more restrained altruistic A1, programmed for developing a large brain. A1s and A5s will both obey the dictate of their limited intra-systemic energy budgets. However, there is an important difference. Genetic Similarity Theory assumes genetic similarity

for altruistic gene(s), as well as independent selection for genes for T-nexus traits. The GTC model sees altruistic acts as based primarily on hormonal similarity. One practical consequences of this is Lo-T A1s generally find Hi-T A5s too loud and physically dominant to share with, irrespective of race, whereas Hi-T A5s find Lo-T A1s too quiet, distant and formally dominant to socialize with. The tentative positioning of races in Table 4 is not intended to exclude that A1-A5 T-altruism contrasts apply cross-racially: A1 Orientals and Blacks may thus find A5 Blacks and Orientals too noisy, whereas A5 Blacks and Orientals may see A1 Blacks and Orientals as too shy, reserved, and difficult to read. A1 blacks and A5 Orientals are just not so common.

More generally, GCO theory maintains that it took at least 40,000 years to single out Northern A1s with large and smart enough brain to survive in cold areas. Today, with the invention of fast, efficient air transportation, it takes less than ten hours to cross many geo-climatic zones without having to pay any cold selection price. Southern A4-5s have recently migrated in numbers to northern A1-2-areas, and now gradually remodel the infra-structures of the societies shaped by A1-2s sifted through unforgiving Darwinian pressures (Nyborg, 2012a). This creates several problems. It has proven difficult to assimilate southern Hi-T/Lo-IQ A4-5s in the modern high-tech northern societies of A1s and A2s. The high heritability for IQ ($\approx$ 0.8) and moderate heritability for T ($\approx$ 0.5) reduces the hope that A5s will be able to care for themselves, their families, and become a valuable asset in these complex societies, despite some plasticity. These problems might feed counterproductive racist attitudes on both sides. Even worse, democracy requires national average IQs above 90 (Vanhanen, 2009).

Also troubling is that European and North American A1-2s have long suffered fertility rates far below the replacement rate of 2.05, whereas non-Western A4-5 migrants generally display higher rates. A1-2 countries will thus according to projection from 1979-2009 data be dominated by A4-5s before the end of this century, a process reinforced by further immigration (Nyborg, 2012a). It

remains to be seen whether this encounter between predominantly altruistic A1-2s and more aggressive A4-5s will damage general altruism and threatens social cohesion in Western societies, and increase psychopathology rates.

Stringer (2011) discussed the emergence of Homo sapiens in Africa, their arrival in Europe and the subsequent extinction of the Neanderthals who migrated from Africa thousands of years before. The GTC model would see the demise of Neanderthals in terms of them becoming too aggressive and "expensive" for cold areas, compared to more recent lightweight Cromagnon cousins. The energy allocated to covariant T-nexus related traits like aggression, strong muscles and heavy bones, drained energy away from IQ-nexus traits, and made them challenge the economy principle, even if they were equally gifted, or perhaps even smarter, than Cromagnons.

Conclusions

Philippe Rushton got his racial rank-ordering of 60+ vital traits for three races right. He did not imply evolutionary progression directly, but GCO theory suggests that A1s and A2s are, in fact, products of such an evolutionary progression.

The flipsides of this evolutionary progression may be that it combines low T and aggression with reduced libido and below replacement fertility. Perhaps, the evolution of northern A1-2s went one bridge too far and exceeded the optimum in the GTC model: Europeans and North American A1-2s seem unwilling - or unfit - to defend their northern territories against more fertile immigrants, thus implicitly consenting to let A4-5s eventually take over their reduced Western world (Nyborg, 2012a).

References

Coleman, D. (2006). Immigration and ethnic change in low-fertility countries: A third demographic transition. *Population and Developmental Review, 32(3),* 401-446.
Coleman, D. (2010). Projections of the Ethnic Minority Populations of the United Kingdom 2006-2056. *Population and Development Review, 36(3),* 441-486.
Dahlstrom et al. (1972). *An MMPI Handbook* (Revised Edition). Vol. I, Clinical Interpretation. Minneapolis: University of Minnesota.
Ellis, L. & Nyborg, H. (1992). Racial/ethnic variations in male testosterone levels: A probable contributor to group differences in health. *Steroids, 57*, 72-75
Gentry, T., Wakefield, J., & Friedman, A. (1985). MMPI Scales for Measuring Eysenck's Personality Factors. *Journal of personality assessment, 49*, 146-149.
Hamilton, W. D. (1964). The genetical evolution of social behavior: I and II. *Journal of Theoretical Biology, 7,* 1-52.
Hoekstra. R.A., Bartels, M., & Boomsma, D.I. (2006). Heritability of testosterone levels in 12-year-old twins and its relation to pubertal development. *Twin Research and Human Genetics, vol. 9,* number 4, 558-565.
Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*, 1-123.
Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.
Jensen, A.R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
Kanazava, S. (2008). Temperature and evolutionary novelty as forces behind the evolution of general intelligence. *Intelligence, 36*, 99-108.
Lovejoy, C. O. (1981). *The origin of man*. Science, 211, 341-350.

Lynn, R. (1991). The evolution of racial differences in intelligence. *The Mankind Quarterly, 32*, 99-121.
Lynn, R. (1977). The intelligence of the Japanese. *Bulletin of the British Psychological Society, 30*, 69–72.
Lynn, R (2006). *Race differences in intelligence: An evolutionary analysis.* Augusta, GA: Washington Summit Publishers.
Lynn, R., & Vanhanen, T. (2006). *IQ and the wealth of nations*. Westport, CT. Praeger.
Nyborg, H. (1987). Covariant trait development across species, races and within individuals. Paper presented in symposium 6: Biology/Genetics, at the Third meeting of the International Society for the Study of Individual Differences, Toronto, Canada, June 18-22nd, Abstract p. 29.
Nyborg, H. (1994). *Hormones, sex, and society: The science of physicology*. Westport, CT: Praeger.
Nyborg, H. (2012a). The decay of Western civilization: Double reversed Darwinian selection. *Personality and Individual Differences, 53,2,* 118-125.
Nyborg, H. (2012b). In conversation with J. Philippe Rushton. (in this book).
Nyborg, H. The suicide of the European Peoples (In prep.).
Nyborg, H. & Jensen, A. R. (2000a). Testosterone Levels as Modifiers of Psychometric g. *Personality and Individual Differences, 28*, 601-607.
Nyborg, H. & Jensen, A. R. (2000b). The Black-White Differences on various psychometric tests: Spearman's hypothesis tested on American Armed services veterans. *Personality and Individual Differences, 28*, 593-599
Orwoll, E., Lambert, L.C., Marshall, L.M., Phipps, K., Blank, J., Barrett-Connor, E., Cauley, J., Ensrud, K., & Cummings, S. (2006). Testosterone and Estradiol among older men. *The Journal of Clinical Endocrinology & Metabolism, 91(4),* 1336-1344.
Ross, R., Bernstein, L., Judd, H., Hanisch, R., Pike, M., & Henderson, B. (1986). Serum testosterone levels in healthy young black and white men. *J. Natl. Cancer Inst., 76*, 45-48.

Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective*. New Brunswick, NJ: Transaction.
Rushton, J.P. (2000). *Race, evolution and behavior* (Third edition). Port Huron, MI: Charles Darwin Research Institute.
Rushton, J. P., Bons, T. A., & Hur, Y-M. (2008). The genetics and evolution of a general factor of personality. *Journal of Research in Personality, 42*, 1173-85. (Corrigendum, 2009, *Journal of Research in Personality, 43,* 532).
Rushton, J.P., & Rushton, E.W. (2004). Progressive changes in brain size and musculo-skeletal traits in seven hominoid populations. *Human Evolution, vol. 19*, no 3, 173-196.
Stringer, G. (2011). *Origin of our Species*. London: Allen Lane.
Templer, D. (2012). Rushton: The Great Theoretician and His Contribution to Personality. (Chapter 6 in this book)
Templer, D. I. & Arikawa, H. (2006). Temperature, skin color, per capita income, and IQ: An international perspective. *Intelligence, 34*, 121−139.
Vanhanen, T. (2009). *The limits to democratization.* Augusts, GA: Washington Summit Publishers.
VES; Vietnam Era Study (1989) *Health Status of Vietnam Veterans, Vol. IV*. (Chapter 5, pp. 67 – 94). Centers for Disease Control, Atlanta, GA.
Wilson, E. O. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Harvard University Press.

Chapter 12

Global Behavioral Variation: A Test of Differential-*K*

Gerhard Meisenberg[1], Michael A. Woodley[2]
Ross University School of Medicine
[1]*Ross University, School of Medicine, Picard Estate, Portsmouth, Dominica, West Indies.*
[2]*Department of Psychology, Umeå University, Sweden and Center Leo Apostel for Interdisciplinary Studies, VUB, Belgium.*

Abstract

Despite Rushton's path-breaking work into evolutionary forces affecting life history traits, not many attempts at operationalizing the differential-*K* spectrum at the level of countries or racial groups have been made so far. We report the construction of a "national *K*" factor from country-level behavioral variables. This *K* factor is closely related to country-level intelligence ("*g*"), operationalized by a composite score of IQ and scholastic achievement. We further demonstrate relationships of both *g* and *K* with measures of current environment and hypothesized evolutionary antecedents. Whereas *K* is predicted most powerfully by intelligence, log-transformed GDP (lgGDP) and skin reflectance, *g* is predicted by skin reflectance, lgGDP, cranial capacity, and a measure of evolutionary novelty.

1. Introduction

Among J. P. Rushton's contributions to psychology, his application of *r-K* theory to individual and group differences has been one of the most consequential and, in its application to race differences, the most controversial. The reason for the continuing strength of this theory can be summarized in one word: parsimony. Using a single theoretical construct, Rushton has provided an explanation for the co-occurrence of many seemingly unrelated traits both at the individual and the racial group levels.

r-K theory had been introduced into biology by MacArthur and Wilson (1967) to describe species differences in the allocation of resources to somatic effort, mating effort, and parenting effort. *K* stands for the carrying capacity of the environment, and *r* for the population's maximal growth rate. *K*-selected traits are thought to be favored by natural selection when the environment is stable and the population is close to the carrying capacity of the ecosystem. They include slow maturation, long life span, production of few offspring, and extensive parental care. The opposite, or *r*-selected strategy, is characterized by fast maturation, short life span, early and prolific reproduction, and little parental care. It was Rushton who first applied *r-K* theory to humans, calling it *differential-K*, owing to the fact that whilst all humans are relatively *K*-selected, some individuals and groups are more *K*-selected than others. He and his co-workers described several psychological traits as *K*-selected, including stable emotional attachments that lead to marital stability, the involvement of fathers in child rearing, future-orientation, and the propensity for long-term planning. Early sexual maturity, early reproduction, and a high incidence of externalizing disorders including antisocial personality are considered elements of an *r*-selected strategy (Bogaert & Rushton, 1989; Rushton 1985, 2000). Many of the predictions made by Rushton and co-workers have been confirmed in subsequent work by A.J. Figueredo and colleagues, who have found significant common *genetic* variance amongst measures of personality, subjective wellbeing and a large array of *r-K* related behaviors (e.g. Figueredo, Vásquez, Brumbach

& Schneider, 2004; Figueredo, Vásquez, Brumbach & Schneider, 2007; Figueredo & Rushton, 2009).

Importantly, differential-*K* (*r*-*K* applied to group differences) and life history theory (*r*-*K* applied to individual differences) are developmental as well as evolutionary theories of human behavior. Unstable environments are expected to favor fast (low-*K*) life history traits such as early puberty and a propensity for social deviance, and being raised in a stable family is expected to favor slow life history (high-K) traits (Ellis, Figueredo, Brumbach & Schlomer, 2009).

Rushton insisted on the evolutionary-genetic origin of individual and especially group differences in differential-*K* related traits. Several objections have been raised against his theory. One is that each of the traits can be explained by environmental conditions, without the need to appeal to evolution (e.g., M. Lynn 1989; Rushton 1989). Slower life history traits can, on first sight, be identified as the typical traits of 'socially privileged' middle class people and prosperous countries, with faster life history traits being typical for low-SES people and for deprived societies and communities (Figueredo, Vásquez, Brumbach & Schneider, 2007). This is consistent with the idea that life history theory comes in "weak" and "strong" forms (Figueredo, Vásquez, Brumbach & Schneider, 2007). The weak form assumes that individual life history traits are not closely related mechanistically. For example early age at puberty and the tendency to engage in antisocial behaviors are influenced by different genes and vary independently at the individual level. However, they are genetically correlated at the population level because those populations that have experienced unstable environments in the past have been selected for both traits. Correlations between these traits at the individual level would result from developmental plasticity or "social privilege", with minimal genetic covariance. The strong form postulates the existence of regulatory genes and physiological mechanisms that coordinate the development of multiple related traits. This version of the theory predicts substantial genetic covariance at the individual-differences level, and is supported by

twin studies (e.g. Figueredo, Vásquez, Brumbach & Schneider, 2004; Figueredo, Vásquez, Brumbach & Schneider, 2007; Figueredo & Rushton, 2009).

A second salient critique is Rushton's use of a crude tripartite system (Caucasoid, Mongoloid, Negroid) that is widely perceived as unsatisfactory for the classification of human population structure (e.g. Anderson, 1991). This critique is potentially obviated by using either more fine-grained racial taxonomies (e.g. Rushton, 2010), or by using continuously variable indicators of recent evolutionary history, such as skin reflectance as a measure of the extent to which the ancestors of an extant population have been exposed to non-tropical (low sunlight) climates with the need for heavy clothing. Relatively recent conditions, such as the presence of agriculture, should be considered because there is evidence for substantial changes in allele frequencies on this time scale (Cochran & Harpending, 2009; Meisenberg, 2008).

Another critique is that the concept of life history theory in humans is redundant because variations in life history traits are explained more parsimoniously by varying levels of intelligence. For example, at the individual level the propensity of marrying, which is a "classical" high-K trait in humans, is favored by high intelligence in societies as diverse as the United States (all races, see Meisenberg & Kaul, 2010) and the Caribbean (Meisenberg et al., 2006). Rushton met this objection by including high intelligence among slow life history traits, along with rule following and sexual restraint. Whilst it could be argued that a slow life history requires foresight and long-term planning, it is not clear why intelligence should be less important for mate seeking than for parenting (Woodley, 2011a).

However, diverse measures of *K* (including measures of ecological preferences and two *K*-batteries including the Arizona Life History Battery and the 'Trifecta') do not usually correlate with *g* at the individual differences scale. A recent meta-analysis involving 12 effect sizes from ten studies found that the two correlate non-significantly (ρ=.023, N=2056), and that the

correlations exhibited significant heterogeneity (Woodley, 2011a). Woodley proposes an alternative conceptualization of the relationship between intelligence and *K* in the form of the Cognitive Differentiation-Integration Effort (CD-IE) hypothesis, which posits that whilst there is no main effect of K on *g*, *K* might affect the relationships among cognitive abilities hierarchically subordinate to *g*: those with high-*K* allocate effort to the cultivation of specialized abilities, whereas those with low-*K* will allocate effort into the development of a generalized ability profile. High-*K* cognitive specialists are capable of exploiting narrow social niches in stable and crowded environments, whereas low-K cognitive generalists acquire domain general skills that can be transferred between broader social niches as a buffer against environmental instability. Furthermore, cognitive generalism permits the creation of more 'multidimensional' mental fitness indicators to aid in short term mating.

CD-IE posits a strong variant of life history theory for explaining the covariance amongst personality, subjective wellbeing and other behavioral indicators of *K*, arguing that directional selection, historically operating on various polymorphisms, is sufficient to account for the existence of strong genetic correlations amongst these variables (Figueredo & Rushton, 2009). Furthermore it predicts the existence of genetic correlations between *K* and changes in the strength of the correlation amongst cognitive abilities. In line with Penke, Denissen and Miller (2007), however, it argues that there are no genetic correlations of non-cognitive life history indicators with *g*, as variability in *g* is maintained within populations mainly via mutation-selection balance. It is posited therefore that *g* corresponds to general neural plasticity, efficiency and speed, which along with other indicators of general fitness, such as fluctuating asymmetry (Miller, 2000), constitutes a source of variance in individual differences genetically independent of *K*.

The present study is a preliminary investigation of the nature of *K* at the country level, including its operationalization via traits that have been included in the life history continuum by

Rushton, and especially in his work on race differences (1985, 1989, 2000, 2004). Our focus is on advancing a more nuanced approach than that taken by Rushton, and in so doing special attention is given to the relationship of "national *K*" with "national *g*".

2. Methods

2.1. Elements of the K factor

A national *K* factor was computed from six life history indicators (for sources, see web references):

1. *Teenage childbearing* is the proportion of children born to mothers aged 19 and below. Data are from the *Demographic Yearbook of the United Nations* (2008). Missing data points were extrapolated from World Bank data.
2. *Contraceptive prevalence* among married couples is averaged from several sources including the *United Nations' Human Development Report* (2004) and the UN Statistics Division.
3. *Sexually transmitted diseases* (*STDs*), from the World Health Report of the WHO (2004 edition), include syphilis, gonorrhea and chlamydia. HIV/AIDS is not included because of its recent African origin, which affects its present geographical distribution.
4. *Homicide rate* (last available date) is from the UN Office of Drugs and Crime.
5. *Crime* is a measure of crime victimization derived from the Gallup World Poll. It is the unrotated first principal component of the proportion reporting theft during the last year, proportion reporting assault/mugging, and proportion feeling unsafe on the streets at night.
6. *Savings rate* is gross domestic savings, 1975-2005 average, from the World Bank.

These variables were selected to capture the elements of early reproduction (teenage pregnancy), promiscuity (STDs), crime

(homicide rates, crime victimization), and future orientation (contraception, savings).

2.2. Development indicators

Intelligence is averaged from two variables: (1) national IQs reported in Lynn & Vanhanen (2006), with amendments and extensions reported in Lynn (2010). Minor corrections were used for Morocco (Sellami et al., 2010) and Saudi Arabia (Batterjee, 2011) based on more recent results; (2) results of international scholastic assessments in mathematics, science and/or reading as described in Meisenberg and Lynn (2011), with additional data points obtained from the International Mathematics Olympiads (Rindermann, 2011).

IQ and school achievement were averaged with weighting for data quality as described in Meisenberg & Lynn (2011). IQ is available for 137 countries and school achievement for 148. IQ correlates with school achievement at r =.875 for the 108 countries having both measures. For countries having only school assessment scores or only IQ, the available measure was used.

lgGDP is log-transformed per capita GDP adjusted for purchasing power from the Penn World Tables 3.6 (Heston et al., 2011), average 1985-2005.

Education measures length of schooling for adults 25+ years old, based on the Barro-Lee data set for 143 countries. Missing data points were extrapolated from World Bank and United Nations sources.

2.3. Evolutionary measures

1. Skin reflectance (skin color), as reported in Jablonski and Chaplin (2000). High values indicate light skin. This is an approximate measure of the proportion of the population's ancestors living in tropical and subtropical regions with high sunlight levels.
2. Evolutionary novelty, measured as distance from sub-Saharan Africa (defined here as Africa south of 20th degree northern latitude), under the assumption that early humans spread along

the coasts of the Indian Ocean before entering North-East Asia and that Europe was settled from south-east. Evolutionary novelty is hypothesized to favor the evolution of higher intelligence, based on the hypothesis that intelligence is an adaptation to conditions that were different from the environment of evolutionary adaptedness (Kanazawa, 2008).

3. Time since the origin of agriculture, as reported by Puttermann and Trainor (2006). Agriculture in seasonal climates is hypothesized to select for foresight and long-term planning.
4. Average winter temperature (January temperature in the northern hemisphere), obtained from Weatherbase. Cold winters have been postulated as an environmental challenge that selected for higher intelligence and cultural complexity (Lynn, 1987).
5. Historical importance of pastoralism in the country, estimated as the percentage of the population depending on this mode of subsistence. Pastoralism may have imposed different selective pressures from settled agriculture.
6. Cranial capacity obtained from Beals, Smith and Dodd (1984). Cranial capacity correlates highly with brain size, which in turn correlates at approximately r=.40 with psychometric intelligence (Rushton & Ankney, 2009).

Six racial groups were defined for the purpose of this study: European, Middle Eastern (West Asia, North Africa, Indian subcontinent), East Asian, Southeast Asian (including Pacific Islanders), and Native American.

3. Results

3.1 Properties of K

A national *K* factor was extracted as the unrotated first factor of a maximum-likelihood factor analysis of the six indicators described under Methods. The correlations of this national *K* factor with its indicators as well as with intelligence (*g*) and log-transformed GDP are shown in Table 1. Most striking is the high correlation of .877

between *K* and *g*. This is about as high as the correlation between school achievement and IQ, the two variables from which *g* was averaged. The close relationship between *K* and *g* is also shown when the correlations of the six K indicators with *K* are correlated with their correlations with *g* (r=.831, N=6, p = .041). At the level of world regions *g* and national K vary in parallel. Average values for *K* range from 69.3 in sub-Saharan Africa to 101.5 in East Asia, and *g* varies from 70.7 in Africa to 103.0 in East Asia.

Table 1. Correlations of K with g, log-transformed GDP, and the K indicators: teenage childbearing (Teenpreg), log of sexually transmitted diseases (lgSTD), contraceptive prevalence, log-transformed homicide rate, crime victimization, and national savings rate. N = 97 countries. Correlations higher than .200 are significant at p<.05.

	K	*g*	lgGDP	Teenpreg	STDs	Contrac.	lgHomic.	Crime
G	.877	1						
lgGDP	-.819	.756	1					
Teenpreg	-.865	-.689	-.665	1				
STDs	-.866	-.877	-.795	.607	1			
Contrac.	.690	.734	.649	-.434	-.774	1		
lgHomic.	-.808	-.638	-.564	.669	.551	-.336	1	
Crime	-.713	-.507	-.463	.575	.454	-.299	.692	1
Savings	.422	.338	.602	-.398	-.304	.192	-.295	-.235

3.2 Evolutionary indicators

Rushton attributed country-level differences in *g* and *K* to genetics and race. Therefore we should first verify whether national intelligence and our national *K* factor do indeed vary systematically with the racial composition of the population. Regression models were constructed in which either *g* or *K* was predicted only by the composition of racial groups residing in the country, with "European" as the omitted control. Regressions were done separately for Old World and New World countries because the racial origins of New World populations can only be estimated. Also, because in the New World countries multiple racial groups coexist or are hybridized, and country-level environmental influences are expected to be important determinants of both *g* and

K, we expect smaller race differences in the New World than the Old World.

Table 2 shows the results, with values for *g* and *K* scaled to 100 for "European." In the New World, the numerically small racial groups other than European, Native American and African were lumped with the Europeans. The adjusted R^2 values show that as expected, racial composition of the population is a less potent predictor of both outcomes in the New World. In the Old World, however, 80% of the variance in *K* and 85% of the variance in *g* is explained by racial denomination. The results suggest that populations with high proportions of Native Americans score slightly lower on *K* than populations with equally high proportions of Africans, although this ranking is reversed for *g*. Otherwise, *g* and *K* generally vary in parallel.

Table 2. Relative scores on g and K for racial groups worldwide, with European scaled to 100. The country-level regressions used racial denomination as the predictor of the outcome measures.

	***K* predicted**			***g* predicted**		
	All countries	Old World	New World	All countries	Old World	New World
European	100.0	100.0	100.0	100.0	100.0	100.0
Middle Eastern	89.4	88.9	-	86.5	86.2	-
African	74.2	72.2	95.4	73.6	73.7	71.0
East Asian	104.7	104.2	-	106.4	106.3	-
Southeast Asian	86.3	85.9	-	90.2	90.0	-
Native American	70.7	-	87.8	77.0	-	75.8
N (countries)	161	128	33	167	136	31
Adj. R^2	.731	.800	.092	.821	.850	.585

If population-level variations in *g* and *K* are based on genetics, they were most likely selected by conditions to which human ancestors were exposed during the last 50,000 years, which is the approximate time scale for modern human racial evolution

(Goebel, 2007); and if they are non-genetic responses to contemporary life conditions, they should be related most closely to indicators of contemporary environmental quality, rather than to past conditions. Table 3 shows correlations of *g* and *K* with plausible evolutionary factors that might have selected for or against these traits in the past. Log-transformed per-capita GDP and length of schooling are included as measures of contemporary environment. We observe again that the correlates of K and *g* are similar. The environmental indicators (lgGDP and education) correlate somewhat more with *K* than with *g*, while the opposite is seen with most of the evolutionary indicators, except for time since agriculture and pastoralism.

Table 3. Correlations of K and g with environmental and evolutionary determinants. Cran. = cranial capacity, Nov. = evolutionary novelty (distance from Africa), Temp. = winter temperature, Agric. = time since introduction of agriculture, and Past. = importance of pastoralism in prehistoric and early historic times. N = 143 countries. For Correlations higher than .165, p<.05.

	K	*g*	Educ.	lgGDP	Skin	Cran.	Nov.	Temp.	Agric.
G	.867	1							
Educ.	.758	.750	1						
lgGDP	.803	.732	.762	1					
Skin	.821	.859	.701	.685	1				
Cran.	.721	.773	.584	.556	.732	1			
Nov.	.273	.433	.337	.319	.485	.461	1		
Temp.	-.705	-.746	-.666	-.484	-.785	-.785	-.215	1	
Agric.	.567	.512	.308	.399	.645	.411	.181	-.419	1
Past.	.051	-.132	-.177	.035	-.079	.063	-.214	.004	.170

This analysis is refined in the regression models of Table 3. Model 1 shows that *g* is predicted by skin reflectance, lgGDP, cranial capacity, evolutionary novelty, and absence of pastoralism. Owing to relatively high colinearity with variance inflation factors (VIFs) up to 8.8), Model 1 was refined by eliminating non-predictors, producing Model 2, in which the highest VIF is 3.4. The

same strategy was followed for the prediction of *K* in Models 3 to 6.

Table 4. Prediction of g and K with development and evolutionary indicators. For definitions of variables, see Table 3.

	***g* predicted**			***K* predicted**		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Nat. intel.					.301**	.315***
lgGDP	.203**	.209***	.332***	.346***	.259***	.281***
Education	.106	.071	.190**	.173**	.108	.080
Skin color	.301**	.430***	.311*	.293***	.276**	.236**
Cranium	.160*	.210***	.100	.103*	.109	.082
Novelty	.250***	.160***	.096	.093	.007	
Temperature	-.054		.015		.066	
Agriculture	.040		.110*	.119**	.083	.090*
Pastoralism	-.095**	-.103**	.031	.071*	.067	
N (countries)	121	124	124	124	113	114
Adj. R^2	.879	.867	.886	.887	.908	.909

* $p < .05$
** $p < .01$
*** $p < .001$

When *g* is excluded from the model, lgGDP is the most important predictor of K followed by skin reflectance, education, and early introduction of agriculture. These models validate Table 3, indicating that evolutionary history is more important for *g*, and contemporary environment is relatively more important for *K*. When *g* is included as a predictor, it becomes a major predictor of *K*, together with lgGDP and skin reflectance (Models 5 and 6).

Attempts to model national *K* as a latent variable when *g* and lgGDP are included were unsuccessful. The correlations of the latter two variables with latent *K* approached unity, and multiple effects of *g* and lgGDP on the *K* indicators were required for acceptable model fit. However, using *K* as a measured variable produced meaningful relations with hypothesized predictors.

In path models predicting *g* without *K*, the important predictors were skin reflectance (β = .464), lgGDP (β = .243), cranial capacity (β = .233), evolutionary novelty (β = .178), and

pastoralism (β = -.118) ($p<.001$, N = 123 countries). Cranial capacity was predicted by winter temperature (β = -.667), evolutionary novelty (β = .356), lgGDP (β = .135), and pastoralism (β = .116). When *K* was predicted with the same variables, significant ($p<.05$) effects were observed for lgGDP (β = .441), skin reflectance (β = .339), evolutionary novelty (β = .129) and time since agriculture (β = .098).

Figure 1 shows a path model including both *g* and *K*. It shows three major effects on *K*: *g*, lgGDP, and skin reflectance (lighter skin → higher *K*). Thus at least one of the evolutionary indicators apparently directly affects *K* independently of *g* and lgGDP.

4. Discussion

Rushton argued that *g* is related to a "slow" (high-*K*) life history, and was differentially selected along with other *K*-related traits in different climatic and ecologic zones. This implies that a close relationship between these two constructs is expected at the species level in animals, and at both the racial group and individual differences level in humans (1985, 2000, 2004).

Evidence indicates an association between *g* and *K* at the racial group differences level (Rushton, 2000), and at the cross-species level when proxies for intelligence such as encephalization quotient are employed (Rushton, 2004). It also needs to be noted that our national *K* factor is not the first attempt at creating a cross-national measure of differential-*K*. Templer (2008) found that measures of infant mortality, GDP, skin reflectance, longevity, total fertility and HIV/AIDS prevalence all share a common source of variance stemming from what he termed a *K* superfactor. Templer

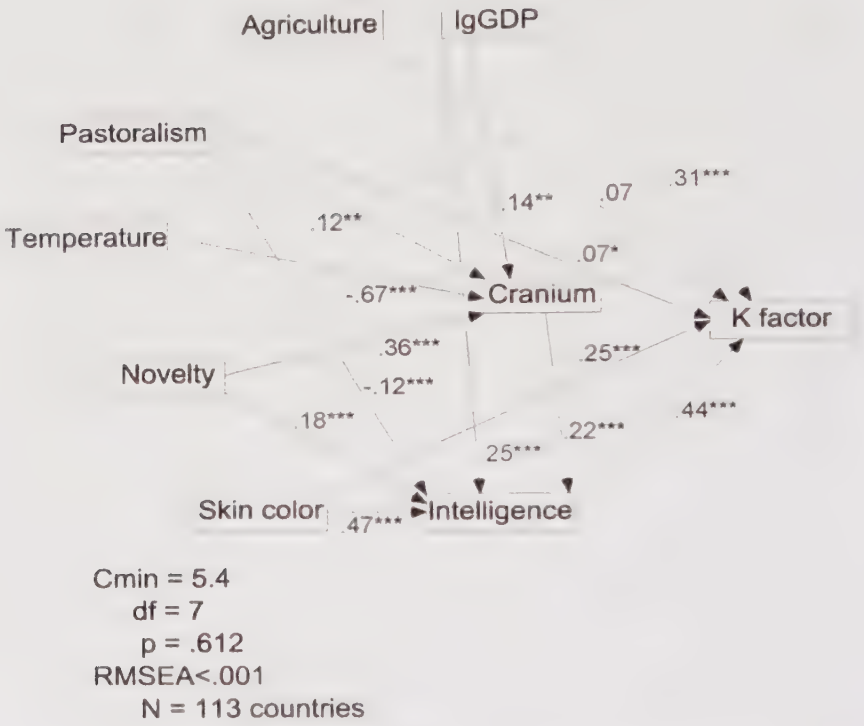


Figure 1. Path model prediction g (intelligence) and K with indicators of evolutionary history and current prosperity. Cranium=cranial capacity.

also included a measure of national IQ amongst his *K* superfactor components, and subsequent work expanded it to include homicide rates (Rushton & Templer, 2009; Templer & Rushton, 2011). We excluded GDP, infant mortality, birth rate and fertility from our definition of *K* because these "development indicators" are known to correlate highly with IQ at the country level (Lynn & Vanhanen, 2006).

However, meta-analysis does not support an association between *g* and *K* at the individual differences level (Woodley, 2011a). What then could account for this "Rushton paradox"?

One possibility is that at the individual differences level, *g* genes assort independently with *K* genes, and became correlated at

the inter-population levels because selective pressures favoring higher *g* also favored higher *K*. A second possibility is population stratification. As Sub-Saharan Africans exhibit low *g* relative to other populations, any trait that systematically differs between Africans and non-Africans would therefore become incidentally correlated with *g* at the population level. A third possibility is that because both *K* and *g* are, in part, dependent upon environmental conditions, poor environments favoring fast life history might also inhibit *g* (Figueredo, 2009). A fourth possibility is that at the level of international comparisons, IQ tests lose their *g*-loadings, such that national differences become accentuated by the Flynn effect. This suggests that effort allocation into the development of specific abilities rather than differences in levels of *g* may partially account for the association between national intelligence measures and national *K* (Woodley, 2011b, in press).

The first question addressed in this research is whether a *K* factor exists. The evidence indicates that something resembling this factor can be isolated from theoretically meaningful measures, and that it exhibits a very high correlation with national intelligence, however it is recognized that the construct in question needs fuller validation. In the path model of Figure 1 this is revealed by a strong direct path between *g* and *K* (β=.44). The path from lgGDP seems to be a little stronger to *K* (.31) than to *g* (.25), suggesting a slightly stronger developmental input into *K*. Skin reflectance also appears to associate more strongly with *g* than with *K* (.47 vs. .25). A history of pastoralism seems to have had opposing effects on *K* and *g*, with the former association being positive (.07) and that latter negative (-.12). It is not clear however whether pastoralism failed to select for higher intelligence in the past, or whether it is better considered a recent cultural or environmental influence. There are also variables which uniquely associate with each construct, evidencing their distinctness. *g* receives a significant path from cranial capacity (.22), which is in turn influenced by temperature (-.67). Contra Rushton (2010), neither of these variables appear to influence *K*. Similarly, evolutionary novelty is associated with *g* both directly and through cranial capacity, independently of *K*.

The results show that *K* can be operationalized at the country level with variables that conform to current understanding of this construct. However, the relationship between country-level *g* and country-level *K*, and especially between country-level *K* and individual-level *K*, require further investigation. Evolutionary conditions appear to be plausible contributors to the current worldwide distributions of both traits.

References

Anderson, J. L. (1991). Rushton's racial comparisons: an ecological critique of theory and method. *Canadian Psychology, 32,* 51-62.

Batterjee, A. A. (2011). Intelligence and education: the Saudi case. *Mankind Quarterly, 52,* 133-190.

Beals, K. L., Smith, C. L., & Dodd, S. M. (1984). Brain size, cranial morphology, climate, and time machines. *Current Anthropology, 25,* 301-330.

Bogaert, A. F. & Rushton, J. P. (1989). Sexuality, delinquency, and *r/K* reproductive strategies: data from a Canadian university sample. *Personality & Individual Differences, 10*, 1071-1077.

Cochran, G. & Harpending, H. (2009). *The 10,000 Year Explosion. How Civilization Accelerated Human Evolution.* New York: Basic Books.

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk: the impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Human Nature, 20*, 204-298.

Figueredo, A. J. (2009). Human capital, economic development, and evolution: a review and critical comparison of Lynn & Vanhanen (2006), and Clark (2007). *Human Ethology Bulletin, 24,* 5–8.

Figueredo, A. J., & Rushton, J. P. (2009). Evidence for shared genetic dominance between the general factor of personality, mental and physical health, and life history traits. *Twin Research & Human Genetics, 12*, 555-563.

Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. R. (2004). The heritability of life history strategy: the *K*-factor, covitality, and personality. *Social Biology, 51,* 121–143.

Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. R. (2007). The *K*-factor, covitality, and personality: a psychometric test of life history theory. *Human Nature, 8,* 47–73.

Goebel, T. (2007). The missing years for modern humans. *Science, 315,* 194-196.
Heston, A., Summers R., & Aten, B. (2011). *Penn World Table version 3.6.* Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania 2009, accessed July15,2011at http://pwt.econ.upenn.edu/php_site/pwt63/pwt63_retrieve.php.
Jablonski, N. G., & G. Chaplin, G. (2000). The evolution of human skin coloration. *Journal of Human Evolution, 39,* 57-106.
Kanazawa, S. (2008). Temperature and evolutionary novelty as forces behind the evolution of general intelligence. *Intelligence, 36,* 99-108.
Lynn, M. (1989). Race differences in sexual behavior: a critique of Rushton & Bogaert's evolutionary hypothesis. *Journal of Research in Personality, 23,* 1-6.
Lynn, R. (1987). The intelligence of the Mongoloids: a psychometric, evolutionary and neurological theory. *Personality and Individual Differences, 8,* 813-844.
Lynn, R. (2010). National IQs updated for 41 nations. *Mankind Quarterly, 50,* 275-296.
Lynn, R., & Vanhanen, T. (2006). *IQ and Global Inequality.* Augusta GA: Washington Summit.
MacArthur, R. H. & Wilson, E. O. (1967). *The Theory of Island Biogeography.* Princeton (NJ): Princeton University Press.
Meisenberg, G. (2008). On the time scale of human evolution: evidence for recent adaptive evolution. *Mankind Quarterly, 48,* 407-443.
Meisenberg, G., Lawless, E., Lambert, E. & Newton, A. (2006). The social ecology of intelligence on a Caribbean island. *Mankind Quarterly, 46,* 395-433.
Meisenberg, G. & Kaul, A. (2010). Effects of sex, race, ethnicity and marital status on the relationship between intelligence and fertility. *Mankind Quarterly, 50,* 151-187.
Meisenberg, G. & Lynn, R. (2011). Intelligence: a measure of human capital in nations. *Journal of Social Political and Economic Studies, 36,* 421-454.

Miller, G. (2000). Sexual selection for indicators of intelligence. In: G.R. Bock, J.A. Goode and K. Webb (eds): *The nature of intelligence: Novartis Foundation symposium 233*, pp. 260-275. Novartis Foundation.

Penke, L., Denissen, J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality, 21,* 549–587.

Puttermann, L., & Trainor, C. A. (2006). *Agricultural transition year country dataset.*
Available from:
http://www.econ.brown.edu/fac/louis_putterman/agricultural%20transition%20data%20set.pdf

Rindermann, H. (2011). Results of the International Mathematical Olympiad (IMO) as indicators of the intellectual classes' cognitive-ability level. In: *Excellence. Essays in honour of Kurt A. Heller*, edited by A. Ziegler and C. Perleth, pp. 303-321. Berlin: Lit.

Rushton, J. P. (1985). Differential K theory: the sociobiology of individual and group differences. *Personality & Individual Differences, 6*, 441-452.

Rushton, J. P. (1989). The evolution of racial differences: a response to M. Lynn. *Journal of Research in Personality, 23,* 7-20.

Rushton, J. P. (2000). *Race, evolution, and behavior: a life history perspective*, 3rd edition. Port Huron (Mich.): Charles Darwin Research Institute.

Rushton, J. P. (2004). Placing intelligence into an evolutionary framework of how g fits into the *r-K* matrix of life-history traits including longevity. *Intelligence, 32,* 321-328.

Rushton, J. P. (2010). Brain size as an explanation of national differences in IQ, longevity, and other life-history variables. *Personality and Individual Differences, 48*, 97-99.

Rushton, J. P. & C.D. Ankney, C. D. (2009). Whole brain size and general mental ability: a review. *International Journal of Neuroscience, 119,* 692-732.

Rushton, J. P., & Templer, D. I. (2009). National differences in intelligence, crime, income, and skin color. *Intelligence, 37*, 341-346.

Sellami, K., Infanzón, E., Lanzón, T., Díaz, A., & Lynn, R. (2010). Sex differences in means and variance of intelligence: some data from Morocco. *Mankind Quarterly, 51,* 84-92.
Templer, D. I. (2008). Correlational and factor analytic support for Rushton's differential-K life history theory. *Personality and Individual Differences, 45*, 440-444.
Templer, D. I. & Rushton, J. P. (2011). IQ, skin color, crime, HIV/AIDS, and income in 50 U.S. States. *Intelligence, 39*, 437-442.
Woodley, M.A. (2011a). The cognitive differentiation-integration effort hypothesis: a synthesis between the fitness indicator and life history models of human intelligence. *Review of General Psychology,* 15:3, 228-245.
Woodley, M.A. (2011b). Heterosis doesn't cause the Flynn effect: a critical examination of Mingroni (2007). *Psychological Review, 118:4,* 689-693.
Woodley, M.A. (In press). A life history model of the Lynn-Flynn effect. *Personality and Individual Differences*. Doi:10.1016/j.paid.2011.03.028.

Web references

Barro-Lee data set: http://www.barrolee.com/data/dataexp.htm
Demographic Yearbook of the United Nations (2008): http://unstats.un.org/unsd/demographic/products/dyb/dyb2.htm
Gallup World Poll: http://www.gallup.com/poll/world.aspx
UN: Human Development Report (2004): http://hdr.undp.org/en/reports/global/hdr2004
UN Office of Drugs and Crime: http://www.unodc.org/unodc/en/data-and-analysis/homicide.html
UN Statistics Division: http://unstats.un.org
Weatherbase: www.weatherbase.com
World Bank data: http://data.worldbank.org/indicator/
World Health Report (2004): http://www.who.int/whr/2004/en/index.html

Chapter13

A Jensen Effect on Dysgenic Fertility: An Analysis involving the National Longitudinal Survey of Youth

Michael A. Woodley[a,b] , Gerhard Meisenberg[c]
[a]*Independent researcher, UK,* [b]*Current affiliations: Umeå University, Sweden & Center Leo Apostel, VUB, Belgium.*
[c]*Ross University School of Medicine.*

Abstract

In this study we attempt to determine whether dysgenic fertility is associated with the Jensen effect. This is investigated with respect to a US population representative sample of 8110 individuals from the National Longitudinal Survey of Youth for whom there exists complete data on IQ and fertility. In addition to the general sample, we also broke the sample out by race and sex so as to examine whether or not the Jensen effect manifested amongst different sub-populations. The method of correlated vectors revealed significant Jensen effects in five of the seven samples, and in all cases the effect was in a direction indicating that subtests with higher *g*-loadings were associated with larger dysgenic fertility gradients. The magnitude of the difference between Spearman's ρ and Pearson's r was non-significant in all cases, suggesting that biasing factors were minimally influencing the result. This finding suggests that dysgenesis occurs on the 'genetic *g*' at the heart of the Jensen effect nexus, unlike the Flynn effect, which is 'hollow' with respect to *g*. Finally, the finding is discussed in the context of two

converging lines of evidence indicating that genotypic IQ or 'genetic *g*' really has been declining over the last century.

1. Introduction

1.1 The Jensen effect

The Jensen effect is a term first coined by Rushton (1998) to describe Arthur Jensen's finding that the vector (rank of the effect magnitude) of many biologically significant variables typically correlates both positively and significantly with the vector of subtest *g*-loadings. In other words the Jensen effect results from the tendency for subtests exhibiting the highest *g*-loadings to be best at discriminating amongst individual differences in biological variables. Jensen's (1998) method of correlated vectors has its origins in the work of Spearman (1927), who proposed what became known as Spearman's hypothesis, namely that Black-White differences would be most pronounced on tests that are the strongest measures of *g*. Recently Spearman's hypothesis has been rebranded as the Spearman-Jensen hypothesis, owing to Jensen's development of the method of correlated vectors (Rushton, 1998). Consistent with the hypothesis, a number of studies have found large and significant Jensen effects on Black-White differences (see: Rushton & Jensen, 2010 for an overview). The nomological net of biological variables for which Jensen effects are known to exist is very broad, and includes inbreeding depression scores, evoked potentials, brain pH, reaction times, test heritabilities (Jensen, 1998; van Bloois, Geutjes, te Nijenhuis & de Pater, 2009), fluctuating asymmetry (Prokosch, Yeo & Miller, 2005), brain size (Rushton & Ankney, 2009), and sex differences (Nyborg, 2005) amongst others.

Even though it is not without its detractors (e.g. Ashton & Lee, 2005), the existence of the Jensen effect would appear to have significant ramifications. Firstly it effectively falsifies Thomsonite models of the development of intelligence (e.g. Bartholomew, Deary & Lawn, 2009; Thomson, 1916; van der Maas et al., 2006), which argue that *g* arises chiefly from random sampling or

mutualistic reinforcement amongst distinct 'neural elements' or bonds, rather than from the existence of a special quality or 'mental energy', as was first posited by Spearman (1927). This is because the Jensen effect reveals the existence of an apparent nexus amongst diverse biological variables and *g*, which suggests that the factor corresponds to something very fundamental to brain neurophysiology, rather than being a mere statistical regularity (Eysenck, 1987; Rushton & Jensen, 2010). Secondly, it permits the 'hollowness' (i.e. the degree to which differences are most pronounced on non-*g* variance rather than *g* variance) of important phenomena like the Flynn effect to be determined. Debate has raged about whether or not the Flynn effect occurs on *g* (see: Wicherts et al., 2004 and Woodley, 2011a for summaries). The most comprehensive study to date is that of te Nijenhuis (in press), who has determined, based on a comparison of studies from the Netherlands and the US, that the Flynn effect is 'hollow', as its effect magnitude is typically slightly negatively correlated with subtest *g*-loadedness. Other apparently 'hollow' effects include, Spearman's Law of Diminishing Returns (Jensen, 2003) and IQ gains due to the effects of retesting (te Nijenhuis, van Vianen, & van der Flier, 2007).

1.2 Dysgenic fertility

The correlation of IQ and its proxies with fertility has been consistently negative since testing began (Lynn, 2011; van Court & Bean, 1985). The existence of this negative correlation had been intuited by Galton in the 1860's (1869). Furthermore in the modern world, the negative correlation appears to be global in extent, existing both within (Meisenberg, 2008) and between countries (Lynn, 2011; Lynn & Harvey, 2008; Meisenberg, 2009; Reeve, 2009). This dysgenic effect, as it became known, has been a considerable source of debate, not least of all because of the existence of the countervailing Flynn effect, which confounded even the earliest attempts to directly quantify it (Lynn, 2011). Even though there has been much speculation about the likely impact of dysgenics on populations (e.g. Cattell, 1936, 1937, 1991; Nyborg,

2012), until recently there has been no concrete evidence of such an impact. The first substantive evidence has been found using temporal correlation analysis (Woodley, 2012a), in which Western genotypic IQ means were reconstructed from 1455 to 2005 based on the assumption that they trended in the same direction as the status-fertility correlation, i.e. positively until the mid 1800s and then negatively afterwards (Skirbekk, 2008). This reconstructed genotypic IQ measure was found to be the strongest predictor of technological and scientific innovation rates, which trend in a similar manner, with peak per capita innovation rates having occurred in the 1870's, followed by a sharp decline (Huebner, 2005). In the study the Flynn effect was reconstructed for previous centuries based on the idea that its rise was relatively shallow up until the beginning of the 20th century, at which point it rose very sharply (three points per decade) and then stagnated after 2000 (Meisenberg, Lawless, Lambert & Newton, 2005). The Flynn effect estimates correlated very strongly with historical estimates of the growth in GDP (PPP) per capita, and in regression a composite measure of these had a strongly inverse relationship with a combined measure of illiteracy and homicide. These findings were interpreted on the basis that genotypic IQ most strongly corresponds with the 'genetic *g*' at the heart of the nomological net of the Jensen effect. As the Flynn effect gains are 'hollow' with respect to *g*, it is argued that these occur on subfactor specific sources of variance, which are free to vary independently of *g*. The negative impact of homicide and illiteracy on the Flynn effect was taken to indicate a contribution from improved environmental conditions, consistent with the life history model proposed in Woodley (2011a, b; 2012b).

Thus far no study has attempted to directly determine if there is a Jensen effect on dysgenic fertility differentials, however the studies of Meisenberg (2010) and Meisenberg and Kaul (2010) have both found evidence that dysgenic fertility differentials might be associated with the effect. In both studies, dysgenics was investigated with respect to fertility differentials on the various subtests of the Armed Services Vocational Aptitude Battery

(ASVAB), which is the principal cognitive ability battery used in the National Longitudinal Survey of Youth (NLSY). The ASVAB has good criterion validity, and formed the backbone of Herrnstein and Murray's (1994) comprehensive analysis of the social determinants of *g*. In Meisenberg (2010) a sample-wide dysgenesis rate of .8 of a point of genotypic IQ per generation was inferred based on the assumption that regression to the mean was to the sample mean. Meisenberg & Kaul (2010) stratified the NLSY based on race and sex and used race-specific means instead of the sample mean. They found a dysgenesis rate of about .4 points per generation for the White cohort, .8 points for the Black cohort and 1.2 points for the Hispanic cohort. In both studies and across racial groups the gradient of dysgenic fertility was found to be most pronounced amongst females and also on *g* rather than the individual subtests. This finding strongly hints at a Jensen effect on dysgenic fertility, which would be significant as it implies that dysgenics is part of the biological nexus of the Jensen effect, and (in line with the suggestion made in Woodley, 2011a, 2012a, b) occurs on 'genetic *g*'. For the first time, the hypothesis that dysgenic fertility is associated with the Jensen effect will here be tested.

2. Methods

Following the protocol in Meisenberg (2010) data were collected from the NLSY79 for all 8110 cases which had information on fertility and intelligence. Briefly, the subjects were aged 14-22 at the inception of the study in 1979.

2.1. Cognitive ability measures

The ASVAB was administered in 1980. The ASVAB includes five academic tests (science, arithmetic, word knowledge, paragraph comprehension, mathematics knowledge), three vocational tests (auto & shop info, mechanical comprehension, electronics info), and two speeded tests (numerical operations, coding speed). The scores of each subtest were age-adjusted separately for race and

sex. The *g*-factor was extracted from these age-adjusted test scores as the unrotated first factor of a maximum-likelihood factor analysis.

2.2. Dysgenic fertility gradient

For 7439 subjects the self-reported number of children was known for 2008. For those who were not interviewed in 2008, responses from the 2006, 2004, or 2002 interviews were used instead. Thus the sample has completed or nearly completed fertility. To calculate the gradient of dysgenic fertility each subtest was correlated with the total number of children.

2.3. Correlated vectors

Spearman's rank correlation and Pearson's product moment correlation were both used (in SPSS) to determine the correlated vector between the magnitude of test *g*-loading and the magnitude of the dysgenic fertility gradient. The N equals the number of ASVAB subtests (10). Fisher's *r*-to-*z* transformation was used to determine whether the two methods produced significantly different magnitude correlated vectors, which would indicate the existence of biasing factors (Nyborg, 2005). The NLSY data were also broken out along race and sex lines, following the protocol of Meisenberg and Kahl (2010) so as to determine whether the Jensen effect exists within these categories. The sample sizes for each of the sub-populations are as follows: White males N=2039, White females N=2102, Black males N=1224, Black females N=1258, Hispanic males N=617 and Hispanic males N=648.

3. Results

Table 1 reveals significant Jensen effects (where the P value for Spearman's ρ is $<.05$) on dysgenic fertility for White males, White females, Black females and Hispanic males, but not for Black males and Hispanic females. Despite this in all instances the correlated vector trends in the expected direction.

Table 1. ASVAB subtests along with their respective g-loadings and dysgenic fertility gradients for the whole of the NLSY sample (N=8110), White males (N=2039), Black males (N=1224), Black females (N=1258), Hispanic males (N=617) and Hispanic females (N=647).

ASVAB subtest	Subtest g-loading (whole sample)	Dysgenic fertility (whole sample)	Subtest g-loading White males	Dysgenic fertility (White males)	Subtest g-loading White females	Dysgenic fertility (White females)	Subtest g-loading Black males	Dysgenic Fertility Black males)	Subtest g-loading Black females	Dysgenic Fertility Black females)	Subtest g-loading Hisp males)	Dysgenic fertility Hisp males)	Subtest g-loading Hisp females)	Dysgenic fertility Hisp females)
Science	916	- 147	895	000	860	- 062	879	- 015	854	- 196	906	- 165	893	- 302
Arithmetic	857	- 103	845	036	850	- 038	781	- 008	787	- 160	804	- 124	814	- 195
Word knowledge	911	- 192	899	- 005	877	- 094	912	- 012	912	- 202	912	- 172	910	- 322
Paragraph comprehension	846	- 131	850	012	799	- 050	851	- 031	852	- 173	869	- 167	871	- 292
Numerical operations	723	- 086	700	050	647	- 036	700	055	696	- 175	721	- 072	702	- 211
Coding speed	655	- 074	641	051	548	004	639	036	613	- 162	682	- 070	669	- 231
Auto-shop information	727	- 105	645	048	616	- 042	715	045	561	- 130	760	- 106	718	- 202
Math knowledge	836	- 130	833	005	834	- 057	756	- 052	793	- 222	808	- 171	811	- 241
Mechanical comprehension	804	- 097	780	046	740	- 039	732	056	831	- 062	807	- 140	720	- 195
Electronic information	847	- 127	823	017	790	- 056	787	015	658	- 102	862	- 210	783	- 189
Pearson's *r*	806**		992***		758*		619 (ns)		806**		855**		547 (ns)	
Spearman's *ρ*	890**		865**		783**		565*		738**		871**		768**	

* $p < .05$.
** $p < .01$.

A difference in the magnitude of Spearman's ρ (the generally preferred indicator of a Jensen effect) and Pearson's r favoring the latter suggests that outliers or peculiarities associated with either the subtest *g*-loadings or the dysgenic fertility magnitudes might be biasing the results (Nyborg, 2005). The differences are generally small however (with the biggest being .221 in the case of Hispanic females and the smallest being .016 in the case of Hispanic males), and are non-significant in all cases, which suggests that bias is minimal.

4. Discussion

It appears that dysgenic fertility is indeed a Jensen effect. Not only is the effect significant in the majority of cases (five out of the seven analyses – in all cases the effect trends in the expected positive direction) but its effect magnitude is relatively large and is on a par with other variables for which Jensen effects have been recorded (e.g. Jensen, 1998; Rushton & Jensen, 2010). This finding using a large and representative sample of the US population, along with subpopulations and a well validated measure of intelligence, therefore allows us to place dysgenic fertility into the genetic nexus of the Jensen effect.

The White males exhibit the least dysgenic fertility of all groups examined, with fertility differentials being positive in most cases. It must be noted however that the abilities with respect to which the White males are apparently in 'eugenic' fertility are the ones exhibiting lower *g*-loadings. Hence, even for this group the tendency is for fertility to be depressed on highly *g*-loaded abilities. This findings of this study help to contextualize other findings in the literature, which have thus far not been discussed in relation to dysgenesis. Silverman (2010), for example, has found that visual reaction time has increased substantially since Galton first started measuring it between 1884 and 1893. In Galton's measures, reaction time means were .183 seconds for men (N=2522) and .187 for women (N=302). Silverman demonstrates that Galton's measures were typical of his era with reference to reviews of

reaction time studies from 1911 (which did not include Galton's studies), in which reaction times ranged from .151 to .200 seconds, with a median of .192.

Silverman's review of twelve modern (post 1941) reaction time studies reveals a substantially longer mean for both men (.250 seconds) and women (.277 seconds), with a combined sample size of 3836. Silverman concludes that in 11 of the 12 studies and in 19 of the 20 comparisons, as well as in his meta-analysis, the differences in reaction time means between Galton's estimates and modern estimates are statistically significant. Silverman speculates that these findings might result from increased exposure to neurotoxins. An alternative hypothesis is that because dysgenic fertility and reaction times are both part of the genetic nexus of 'genetic *g*', dysgenic fertility is prolonging reaction time in addition to reducing *g*. This would make Silverman's finding the first strong evidence that genotypic IQ ('genetic *g*') really has been in decline since the last decades of the 19th century.

Another issue that our finding may shed light upon is the apparent anti-Flynn effect on measures of crystallized intelligence (in particular crystallized knowledge), which has been observed in a number of countries, sometimes spanning multiple decades (e.g. Khaleefa, Sulman & Lynn, 2008; Lynn, 2009; te Nijenhuis, in press; Wicherts et al., 2004). Declines in measures of crystallized intelligence typically solicit little commentary in the literature, as gains in measures of fluid intelligence usually swamp these losses (e.g. Khaleefa, Sulman & Lynn, 2008; Lynn, 2009; te Nijenhuis, in press; Wicherts et al., 2004). A handful of researchers have proposed that measures of crystallized intelligence might be superior measures of *g* relative to measures of fluid intelligence (Gignac, 2006; Gregory, 1999; Matarazzo, 1972; Robinson, 1999). Consistent with this is the finding that in head-to-head comparisons using a sufficiently broad array of intelligence measures, measures of fluid intelligence (such as Ravens matrices) have been found to be substantively less *g*-loaded than measures of crystallized knowledge (such as vocabulary, see: Johnson, Bouchard Jr, Krueger, McGue & Gottesman, 2004). Furthermore Miller (2000a)

has argued that crystallized knowledge (such as vocabulary) is likely to have been (and still is) a very significant indicator of general fitness (underlying genetic quality), which might therefore account for its relatively high *g*-loading, as Miller has also argued that the positive manifold results from the effects of pleiotropic mutations which in abundance would lower both *g* and general fitness (Miller, 2000b, c). The observations of multi-decadal declines in crystallized knowledge might therefore relate to dysgenic fertility via the genetic nexus as both anti-Flynn effects and dysgenic fertility concern subtests exhibiting high *g*-loadings.

The finding of a Jensen effect on dysgenic fertility is significant to future research, as in light of the aforementioned findings it strongly suggests that 'genetic *g*' has been in decline since the end of the 19th century. This reinforces the significance of the idea that changing genotypic IQ has had real world impacts on important factors such as the rates of scientific and technological innovation amongst Western populations (Woodley, 2012a).

References

Ashton, M. C., & Lee, K. (2005). Problems with the method of correlated vectors. *Intelligence, 33,* 431–444.

Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review, 116,* 567-579.

Cattell, R. B. (1936). Is our national intelligence declining? *Eugenics Review, 28,* 181–203.

Cattell, R. B. (1937). *The fight for our national intelligence.* London: King & Sons.

Cattell, R. B. (1991). Some changes in social life in a community with a falling intelligence quotient. *Mankind Quarterly, 31,* 323-344.

Eysenck, H. J. (1987). Thomson's "bonds" or Spearman's "energy": Sixty years on. *Mankind Quarterly, 27,* 259-274.

Galton, F. (1869). *Hereditary genius.* London: MacMillan.

Gignac, G. E. (2006). Evaluating subtest saturation levels via the single trait-correlated uniqueness (STCU) SEM approach: Evidence in favor of crystallized subtests as the best indicators of '*g*'. *Intelligence, 34,* 29-46.

Gregory, R. J. (1999). *Foundations of intellectual assessment: The WAIS-III and other tests in clinical practice.* Boston: Allyn and Bacon.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press Paperbacks.

Huebner, J. (2005). A possible declining trend for worldwide innovation. *Technological Forecasting and Social Change, 72,* 980–986.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport: Praeger.

Jensen, A. R. (2003). Regularities in Spearman's law of diminishing returns. *Intelligence, 31,* 95–105.

Johnson, W., Bouchard Jr., T. J., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one *g*: Consistent results from three test batteries. *Intelligence, 32,* 95-107.
Khaleefa, O., Sulman, A., & Lynn, R. (2008). An increase of intelligence in Sudan, 1987-2007. *Journal of Biosocial Sciences, 41,* 279-283.
Lynn, R. (2009). Fluid intelligence but not vocabulary has increased in Britain, 1979-2008. *Intelligence, 37*, 249-255.
Lynn, R. (2011). *Dysgenics: Genetic deterioration in modern populations* (Second revised edition). London: Ulster Institute for Social Research.
Lynn, R., & Harvey, J. (2008). The decline of the world's IQ. *Intelligence, 36*, 112–120.
Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). New York: Oxford University Press.
Meisenberg, G. (2008). How universal is the negative correlation between education and fertility? *Journal of Social, Political and Economic Studies, 33,* 205–227.
Meisenberg, G. (2009). Wealth, intelligence, politics and global fertility differentials. *Journal of Biosocial Sciences, 41,* 519–535.
Meisenberg, G. (2010). The reproduction of intelligence. *Intelligence, 38,* 220–230.
Meisenberg, G., & Kaul, A. (2010). Effects of sex, race, ethnicity and marital status on the relationship between intelligence and fertility. *Mankind Quarterly, 50,* 151–187.
Meisenberg, G., Lawless, E., Lambert, E., & Newton, A. (2005). The Flynn effect in the Caribbean: Generational change of cognitive test performance in Dominica. *Mankind Quarterly, 46,* 29–69.
Miller, G. F. (2000a). *The mating mind: How sexual choice shaped the evolution of human nature.* New York: Doubleday.
Miller, G. F. (2000b). Sexual selection for indicators of intelligence. In G. R. Bock, J. A.
Goode, & K. Webb (Eds.), *The nature of intelligence* (pp. 260–275). Chichester: Wiley Ltd.

Miller, G. F. (2000c). Mental traits as fitness indicators: Expanding evolutionary psychology's adaptationism. *Annals of the New York Academy of Sciences, 907,* 62–74.
Nyborg, H. (2005). Sex related differences in general intelligence *g*, brain size and social status. *Personality and Individual Differences, 39,* 497-509.
Nyborg, H. (2012). The decay of Western civilization: Double relaxed Darwinian selection. *Personality and Individual Differences, 53,* 118-125.
Prokosch, M., Yeo, R., & Miller, G. (2005). Intelligence tests with higher *g*-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence, 33,* 203–213.
Reeve, C. L. (2009). Expanding the *g*-nexus: Further evidence regarding the relations among national IQ religiosity and national health outcomes. *Intelligence, 37,* 495-505.
Robinson, D. L. (1999). The 'IQ' factor: Implications for intelligence theory and measurement. *Personality and Individual Differences, 27,* 715–735.
Rushton, J. P. (1998). The "Jensen effect" and the "Spearman-Jensen hypothesis" of Black-White IQ differences. *Intelligence, 26,* 217-225.
Rushton, J. P., & Ankney, C. D. (2009). Whole brain size and general mental ability: A review. *International Journal of Neuroscience, 119,* 691-731.
Rushton, J. P., & Jensen, A. R. (2010). The rise and fall of the Flynn Effect as a reason to expect a narrowing of the Black-White IQ gap. *Intelligence, 38,* 213–219.
Silverman, I. W. (2010). Simple reaction time: It is not what it used to be. *American Journal of Psychology, 123,* 39-50.
Skirbekk, V. (2008). Fertility trends by social status. *Demographic Research, 18,* doi:10.4054/DemRes.2008.18.5
Spearman, C. (1927). *Abilities of man: Their nature and measurement.* New York: Macmillan.

te Nijenhuis, J. (In press). The Flynn effect, group differences and *g* loadings. *Personality and Individual Differences*. doi:10.1016/j.paid.2011.12.023.
te Nijenhuis, J., van Vianen, A. E. M., & van der Flier, H. (2007). Score gains on *g* loaded tests: No *g*. *Intelligence, 35,* 283–300.
Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology, 8,* 271–281.
van Bloois, R. M., Geutjes, L.-L., te Nijenhuis, J., & de Pater, I. E. (2009, December). *g* loadings and their true score correlations with heritability coefficients, giftedness, and mental retardation: Three psychometric meta-analyses. Paper presented at the Symposium on Group Differences, 10th Annual Meeting of the International Society for Intelligence Research, Madrid, Spain.
van Court, M., & Bean, F. D. (1985). Intelligence and fertility in the United States: 1912–1982. *Intelligence, 9,* 23–32.
van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113,* 842–861.
Wicherts, J. M., Dolan, C. V., Hessen, D. J., Oosterveld, P., van Baal, G. C. M., Boomsma, D. I., et al. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence, 32,* 509–537.
Woodley, M. A. (2011a). Heterosis doesn't cause the Flynn effect: A critical examination of Mingroni (2007). *Psychological Review, 118,* 689-693.
Woodley, M. A. (2011b). The cognitive differentiation-integration effort hypothesis: A synthesis between the fitness indicator and life history models of human intelligence. *Review of General Psychology, 15*, 228-245.
Woodley, M. A. (2012a). The social and scientific temporal correlates of genotypic intelligence and the Flynn effect. *Intelligence, 40,* 189-204.
Woodley, M. A. (2012b). A life history model of the Lynn-Flynn effect. *Personality and Individual Differences, 53,* 152-156.

Index

A1-A5 continuum; 321
achievement; 75; 104; 119; 120; 123; 171; 175; 177; 180; 317; 327; 333; 335
adaptation; 67; 70; 152; 156; 167; 195; 200; 213; 222; 227; 238; 249; 252; 272; 287; 290; 291; 315; 317; 318; 334; 345
aggregation; 78; 231; 235; 246
aggression; 213; 217; 227; 234; 240; 246; 271; 290; 291; 305; 307; 309; 315; 316; 317; 322; 324; 325
agriculture; 125; 273; 330; 334; 335; 338; 339; 340
altruism; 59; 66; 215; 217; 233; 234; 235; 237; 238; 239; 242; 244; 245; 246; 255; 257; 263; 267; 268; 269; 270; 271; 272; 274; 278; 280; 282; 285; 286; 287; 317; 318; 320; 322; 324
androgyne; 307
androtype; 308; 321; 322
assortative mating; 200; 205; 214; 271
ASVAB; 352; 353;

Behavior genetics; 49; 73; 77; 248; 249; 252; 270
blacks; 65; 68; 69; 73; 74; 83; 103; 289; 294; 306; 311; 322; 323
Black-White differences; 47; 50; 52; 53; 65; 103; 124; 349
body size; 55; 61; 67
brain size; 47; 55; 56; 58; 60; 61; 63; 65; 69; 72; 73; 74; 76; 77; 78; 82; 85; 99; 117; 124; 216; 217; 288; 299; 321; 329; 335; 347; 349; 362; 363
Burt C; 49; 214

Cavalli-Sforza, L.; 116; 121; 160; 173; 268; 284; 302; 305
243; 265; 306; 329; 347
chromosomes; 73
civilization; 308; 328; 362
climate; 56; 63; 106; 217; 218; 224; 290; 291; 292; 316; 317; 323; 330; 344
Cochran H; 330; 344
cognitive ability; 65; 85; 103; 104; 105; 107; 108; 110; 111; 112; 113; 114; 117; 118; 119; 123; 124; 125; 127; 128; 131; 134;

137; 140; 147; 151; 153; 156; 157; 158; 159; 160; 161; 162; 163; 165; 166; 167; 168; 169; 170; 171; 175; 178; 271; 306; 352
Cognitive Differentiation-Integration Effort (CD-IE) hypothesis; 331
Coleman D; 228; 326
Cosmides, L.; 249; 260; 272; 281; 285; 287
covariation; 247; 252; 259
cranial capacity; 55; 56; 72; 78; 82; 327; 338; 339; 341; 343
crime; 48; 58; 68; 216; 222; 225; 227; 232; 233; 266; 275; 322; 333; 336; 347
Cromagnon; 324
crystallized intelligence; 104; 112; 129; 140; 148; 154; 156; 171; 179; 357; 360
cultural strategies; 271

Darwin, C.; 49; 106; 123; 178; 214; 220; 248; 260; 261; 288
Dawkins, R.; 115; 116; 121; 148; 174; 253; 261
Deary, I.; 169; 174; 349; 360
default hypothesis; 74
democracy; 308; 317; 324
dexterity; 90; 92; 93; 97
Differential K Theory; 64; 213; 214; 215; 218; 222; 224
differential psychology DP; 175; 248; 259
disease; 75; 223; 233; 251
dysgenesis; 348; 352; 356

East Asians; 48; 50; 57; 58; 59; 60; 65; 68; 69; 72; 73; 216; 221; 227; 288; 292; 297; 298; 299
economy principle; 309; 324
Eibl-Eibesfeldt, I.; 268; 269; 270; 284
energy; 67; 68; 128; 215; 307; 315; 316; 317; 318; 322; 323; 324; 349; 360
epigenetic; 264; 271
estradiol; 309; 328
ethnic identity; 269; 282
ethnic nepotism; 215; 239; 271; 272; 274; 280; 281; 282; 284
ethology; 270
europeans; 48; 58; 60; 72; 104; 116; 119; 133; 141; 147; 149; 160; 171; 172; 226; 288; 292; 293; 294; 300; 302; 325; 336
evoked potentials; 51; 349
evolution; 55; 61; 63; 64; 70; 72; 73; 77; 85; 116; 123; 176; 178; 213; 227; 231; 235; 244; 246; 251; 261; 263; 265; 267; 268; 271; 284; 286; 306; 315; 318; 322; 325; 327; 329; 334;

337; 344; 345; 346; 347; 362
evolutionary novelty; 334
evolutionary progression; 307; 308; 321; 325
evolutionary psychology; 66; 76; 248; 262; 287; 362
evolutionary tree; 72
exceptional achievement; 114
extended phenotype; 116; 121; 148; 169; 174
eye color; 226
Eysenck, H.J.; 49; 188; 214; 228; 239; 246; 327; 350; 360

Family structure; 47
fertility; 58; 316; 324; 325; 326; 340; 341; 346; 348; 350; 351; 352; 353; 354; 356; 357; 358; 361; 362; 364
Figueredo, A.J.; 220; 229; 247; 250; 252; 254; 255; 256; 257; 258; 259; 260; 261; 262; 263; 264; 266; 267; 268; 329; 332; 342; 344; 345
fitness; 67; 68; 179; 199; 208; 238; 246; 249; 266; 267; 268; 269; 270; 271; 272; 279; 331; 332; 347; 358; 362; 364
Five Factor Model; 180
fluid intelligence; 101; 112; 126; 127; 129; 134; 135; 139; 140; 147; 148; 156; 171; 179; 357
Flynn effect; 52; 65; 66; 82; 85; 86; 87; 88; 93; 95; 96; 97; 99; 101; 102; 103; 105; 121; 130; 133; 149; 174; 175; 176; 177; 180; 265; 342; 347; 348; 350; 357; 362; 363; 364

g factor; 50; 52; 63; 65; 74; 85; 102; 103; 230; 327; 361
Galton, F.; 49; 61; 62; 180; 209; 214; 220; 248; 264; 350; 356; 357; 360
GATB; 89; 90; 91; 92; 94; 96; 97; 101; 105
GDP; 99; 133; 327; 334; 335; 336; 338; 340; 341; 351
general factor of personality (GFP); 179; 207; 210; 212; 213; 214; 213; 219; 232; 242; 251; 256; 265; 266; 267; 268
general intelligence (*g*); 70; 73
General Trait Covariance model; 307; 308; 318
genetic *g*; 348; 351; 352; 357; 358
genetic similarity theory; 214; 238; 246; 264; 265;

267; 270; 271; 272; 273; 274; 275; 277; 286; 323
genomics; 268; 275
genotypic IQ; 348; 351; 352; 357; 358
geo-climatic changes; 318
geography; 56; 75; 284; 305
gestation; 58; 61; 288
General Factor of Personalty; 179; 180; 181; 182; 184; 189; 190; 192; 193; 194; 195; 196; 197; 198; 199; 200; 201; 204; 205; 210; 213; 219; 220; 221; 222; 227; 232; 234; 235; 242; 243; 251; 256
Gottfredson, L.; 62
group selection; 245; 269; 271; 282; 284; 286
GST; 267; 269; 271; 272; 274; 275; 280; 281; 282
General Trait Covariance; 308; 316; 318; 319; 320; 321; 323; 324; 325

Hair color; 226
Hamilton B; 238; 244; 253; 263; 267; 269; 270; 272; 284; 285; 327
Harpending, H.; 61; 79; 267
heritability; 50; 51; 54; 74; 76; 77; 88; 104; 124; 199; 200; 214; 246; 248; 252; 257; 262; 307; 324; 344; 363
Herrnstein, R.; 222; 229; 352; 360
hormones; 59; 77; 216; 240; 318; 321; 323
Hunt, E.; 104; 117; 118; 122; 147; 159; 176; 186; 212
hunter-gatherer; 269; 282
hybrid vigor; 99

Immigrants; 86; 104; 325
inbreeding depression; 51; 52; 64; 87; 88; 99; 103; 349
individual differences; 49; 54; 115; 119; 213; 232; 237; 238; 246; 247; 248; 250; 251; 253; 263; 265; 312; 329; 331; 332; 340; 342; 349
innovation rates; 351
intelligence; 47; 60; 63; 64; 65; 67; 66; 69; 70; 73; 74; 76; 77; 78; 82; 83; 84; 85; 86; 89; 95; 101; 102; 105; 103; 104; 106; 110; 112; 113; 115; 116; 118; 119; 122; 123; 124; 126; 127; 128; 129; 130; 131; 133; 134; 135; 136; 137; 139; 140; 142; 147; 149; 154; 158; 159; 160; 161; 165; 169; 171; 172; 173; 175; 176; 178; 179; 180; 200; 211; 214; 213; 215; 216; 217; 218; 222; 224; 227; 232; 242; 250; 257; 262;

264; 265; 266; 267; 271; 288; 290; 291; 299; 306; 327; 330; 331; 334; 335; 336; 340; 341; 342; 345; 346; 347; 349; 352; 356; 357; 360; 361; 362; 364
inter-male aggression; 290; 292
intra-systemic; 307; 315; 317; 318; 322; 323
IQ/T-Geo-Climatic Origin (GCO) theory; 307; 308; 315; 318
IQ-Nexus; 307; 308; 316; 321
Irwing, P.; 179; 180; 186; 197; 199; 200; 201; 210; 213; 219; 220; 228; 231; 232; 242; 246; 251; 257; 265

Jensen, A.R.; 47; 49; 50; 51; 52; 53; 55; 56; 57; 63; 64; 65; 66; 67; 68; 74; 85; 86; 87; 88; 99; 101; 102; 103; 104; 117; 118; 121; 124; 158; 176; 214; 219; 222; 225; 230; 310; 316; 327; 328; 348; 349; 351; 354; 355; 356; 358; 361; 363
Jensen effect; 51, 64; 65; 348; 349; 351; 354; 355; 356; 358; 363

K factor; 258; 327; 332; 335; 336; 340; 342
Kanazawa, S.; 290; 305; 334; 345
kin selection; 214; 238; 239; 286
kinship; 267; 268; 269; 270; 272; 273; 274; 275; 276; 277; 278; 279; 282; 317; 323
K-strategists; 58; 59; 321

Latitude; 56; 320; 334
libido; 325
life span; 47; 59; 103; 216; 328
life-history; 57; 60; 61; 65; 67; 70; 71; 72; 74; 75; 266; 347
London School; 49; 66
longevity; 58; 61; 65; 123; 179; 250; 251; 264; 265; 340; 347
Lynn, R.; 57; 130; 152; 159; 287

Magnetic resonance imaging (MRI); 56
mating strategies; 68; 76
maturation; 58; 67; 69; 78; 132; 134; 215; 250; 261; 288; 316; 328
Meisenberg, G.; 105; 123; 133; 152; 159; 165; 170; 177; 218; 231; 327; 330; 333; 345; 346; 350; 352; 353; 361; 362
mental chronometry; 49

method of correlated vectors; 51; 52; 53; 89; 92; 94; 95; 97; 99; 348; 349; 360
meta-analysis; 99; 102; 161; 185; 209; 210; 257; 267; 300; 305; 306; 331; 342; 357
migration; 287; 307; 308; 315; 316; 322
MMPI; 180; 182; 183; 196; 197; 204; 209; 210; 219; 310; 319; 320; 326; 327
mob science; 65; 81; 82
modular; 249
molecular; 315; 316; 323
moral; 76; 119; 223; 235; 236; 245; 283; 287
MTMM; 181; 182; 184; 185; 186; 201; 204; 205
Multitrait-Multimethod; 181; 209
Murray, C.; 114; 122; 222; 229; 261; 352; 360

National IQ; 341; 362
national *K*; 327; 332; 335; 336; 339; 340; 342
National Longitudinal Survey of Youth; 348; 352
neanderthals; 324
negroids; 68; 287; 288; 289; 292; 293; 294; 295; 297; 298; 300; 301; 302; 303
nepotism; 267; 269; 280; 286; 287
nerve conduction velocity; 51; 99
Neural complexity; 60
Nisbett, R.E.; 86; 88; 97; 102; 118; 122
Nobel prizes; 114
north africans; 288; 302
Nyborg, H.; 49; 58; 64; 67; 68; 106; 122; 197; 210; 307; 308; 309; 310; 312; 315; 317; 320; 321; 324; 325; 327; 328; 349; 350; 353; 356; 362

Obituary; 76
offspring; 52; 58; 59; 68; 70; 71; 87; 215; 217; 231; 238; 276; 288; 328
OPQ; 186; 187; 188; 190; 193; 209

Pair bonding; 67; 71; 253
parental care; 58; 215; 238; 328
pastoralism; 335; 338; 340; 342
paternal investment theory; 291
Pearson K; 49; 92; 180; 214; 312; 313; 319; 348; 353; 355
penis size; 287; 288; 289; 291; 292; 293; 294; 295; 298; 299; 300; 301; 302; 303; 304; 305; 307

personality; 47; 58; 61; 66; 69; 78; 179; 180; 181; 182; 183; 184; 185; 186; 187; 195; 196; 198; 201; 204; 205; 206; 207; 208; 209; 210; 211; 212; 213; 214; 215; 213; 214; 215; 219; 220; 221; 223; 224; 226; 227; 228; 229; 231; 232; 234; 235; 237; 239; 242; 244; 245; 246; 247; 249; 250; 251; 253; 256; 257; 260; 262; 264; 265; 267; 271; 286; 309; 310; 318; 320; 327; 329; 328; 331; 344; 345; 346
phenotype; 73; 89; 200; 258; 269; 271; 273; 281; 282
plasma testosterone; 310
polymorphisms; 275; 276; 331
progressive evolution; 72; 73; 106; 304; 309; 315
prosocial behavior; 227; 233; 234; 235; 236; 239; 240; 244; 247; 269
psychopathology; 182; 183; 208; 210; 223; 232; 318; 320; 322; 324

Race differences; 47; 52; 56; 57; 59; 61; 64; 65; 67; 70; 74; 75; 76; 77; 78; 82; 84; 85; 88; 101; 102; 103; 121; 122; 138; 153; 217; 239; 254; 264; 265; 267; 281; 285; 287; 288; 289; 291; 292; 295; 298; 301; 303; 308; 311; 312; 313; 321; 322; 323; 328; 332; 336; 346; 348; 352; 353; 361
racial realism; 214
racism; 77; 168; 169
Raven's Matrices; 53; 65; 66; 86; 88; 96; 358
reaction time; 49; 50; 51; 70; 356; 357; 363
a reduced Western world; 325
religion; 269; 282; 288
reproduction; 58; 68; 69; 71; 126; 215; 248; 250; 260; 263; 264; 306; 316; 328; 333; 344; 361
resolute ignorance; 82
Rindermann, H.; 103; 105; 110; 112; 114; 115; 117; 122; 123; 125; 130; 131; 137; 140; 150; 153; 157; 159; 162; 165; 169; 178; 333; 346
r-K life history theory; 57; 58; 66; 68; 71; 72; 266; 287; 288; 291; 300; 328

SACMEQ; 107; 152; 153; 158; 159; 165; 167; 175; 177
Salter F; 267; 275; 276; 280; 282; 283; 284; 285; 287
SAT; 104
scientific racists; 77

secular gains in IQ; 86; 88; 103
selection theory; 270
SES; 112; 117; 147; 148; 153; 161; 329
sex; 56; 58; 59; 67; 69; 75; 76; 77; 175; 216; 217; 224; 231; 240; 248; 253; 261; 265; 281; 309; 328; 345; 346; 348; 349; 352; 353; 361
skin brightness; 70; 108; 110; 111; 157; 158; 162; 163; 165; 166; 167; 168; 171; 176; 180; 218; 225; 227; 232; 233; 266; 307; 327; 329; 330; 334; 338; 339; 340; 347
Snyderman, M.; 118; 124; 160; 179
social class; 55
social cohesion; 268; 285; 324
social desirability; 180; 181; 182; 183; 186; 187; 194; 204; 206; 208; 209
social learning theory; 233; 235; 237
social organization; 59; 67; 69; 71; 78; 215; 216
social selection; 248
sociobiology; 64; 179; 231; 233; 237; 253; 264; 306; 346
solidarity; 268; 280; 281
south asians; 52; 103; 288; 293; 303
south-north gradient; 317
Spearman, C.; 47; 49; 50; 51; 53; 64; 66; 101; 139; 214; 219; 222; 232; 312; 313; 319; 328; 348; 349; 353; 354; 355; 360; 361; 363
Spearman's hypothesis; 47; 50; 53; 64; 349
survival; 58; 248; 307; 317

te Nijenhuis, J.; 85; 86; 91; 97; 99; 104; 112; 123; 130; 133; 137; 178; 180; 201; 211; 257; 266; 267; 349; 350; 357; 363
temperament; 47; 58; 215
temperature; 125; 218; 334; 338; 340; 343
Templer, D.; 158; 180; 213; 215; 217; 218; 221; 224; 225; 226; 230; 232; 233; 250; 266; 290; 293; 298; 299; 307; 311; 315; 329; 340; 341; 347
testosterone; 217; 225; 228; 232; 240; 244; 290; 291; 292; 305; 306; 307; 309; 311; 313; 314; 316; 319; 327; 328
The "out-of-Africa" theory; 59
TIMSS; 104; 107; 150; 151; 153; 156; 158; 160; 178
T-Nexus; 307; 316; 317; 320

Tooby,. J; 249; 260; 272; 281; 285; 287
trade-off; 323
trait clusters; 249
Trivers, R.L.; 253; 266
twins; 50; 51; 52; 54; 59; 65; 103; 213; 220; 233; 238; 239; 240; 241; 242; 244; 246; 247; 255; 258; 327; 330
two-pillar theory of adaptive functioning; 213

Unacceptable science; 80
universal Darwinian selection; 315; 323

Vanhanen, T.; 57; 64; 86; 102; 104; 108; 110; 122; 127; 130; 131; 140; 150; 157; 158; 165; 177; 218; 230; 275; 287; 312; 317; 324; 328; 329; 333; 341; 344; 345
Vietnam Era Study (VES); 310; 318; 329

Weber, M.; 268; 269; 274; 288
Wechsler Intelligence Scale for Children (WISC); 52
whites; 65; 68; 69; 73; 74; 83; 103; 289; 294; 306; 311; 322
Wicherts, J.; 88; 105; 104; 108; 110; 111; 116; 118; 124; 133; 150; 159; 160; 162; 165; 167; 170; 177; 180; 350; 357; 364
Williams, G.C.; 105; 121; 233; 253; 267
Wilson, E.O.; 57; 64; 66; 186; 208; 213; 215; 230; 249; 253; 263; 267; 282; 288; 307; 329; 328; 345
within-race variation; 70; 74
Woodley, M.; 87; 105; 117; 123; 169; 178; 247; 251; 267; 331; 342; 347; 348; 350; 351; 352; 358; 364

The Life History Approach to Human Differences
A Tribute to J. Philippe Rushton

Edited by Helmuth Nyborg

Published 2015 by Ulster Institute for Social Research, London.

A collection of articles by eminent scientists, first published in *Personality and Individual Differences* to mark the death in 2012 of J. Philippe Rushton. His ground-breaking and controversial research in the fields of evolutionary theory, genetic similarity theory, *r/K*-theory, life history theory, sex differences, brain size differences, and altruism was the catalyst for fierce debate within both the academic community and the wider population. The debate is one that continues to this day.

Contributors to the thirteen papers include the late Arthur Jensen, Heiner Rindermann, Richard Lynn and Gerhard Meisenberg. The book includes an obituary to J. Philippe Rushton and a conversation between Rushton and Helmuth Nyborg.

Front cover photo of J. Philippe Rushton courtesy of Stephen Rushton.

Available in paperback and electronic formats from *www.ulsterinstitute.org*
Price £20 (paperback) £5 (ebook)

ISBN 978-0-9930001-1-9